Unlocking the Use of Raw Multispectral Earth Observation Imagery for Onboard Artificial Intelligence

Gabriele Meoni [ID], Roberto Del Prete, Federico Serva, Alix De Beusscher, Olivier Colin, and Nicolas Longépé [ID]

Abstract—Nowadays, there is growing interest in applying artificial intelligence (AI) on board Earth observation (EO) satellites for time-critical applications, such as natural disaster response. However, the unavailability of raw satellite data currently hinders research of lightweight preprocessing techniques and limits the exploration of end-to-end pipelines, which extract insights directly from the source data. To fill this gap, this work presents a novel methodology to automate the creation of datasets for the detection of target events (e.g., warm thermal hotspots) or objects (e.g., vessels) from Sentinel-2 raw data and other multispectral EO pushbroom raw imagery. The presented approach first processes the raw data by applying a pipeline consisting of a spatial band registration and georeferencing of the raw data pixels. Then, it detects the target events by leveraging event-specific state-of-the-art algorithms on the Level-1 C products, which are mosaicked and cropped on the georeferenced correspondent raw granule area. The detected events are, finally, reprojected back on the corresponding raw images. We apply the proposed methodology to realize thermal hotspots in raw Sentinel-2 data (THRawS), the first dataset of Sentinel-2 raw data containing warm thermal hotspots. THRawS includes 1090 samples containing wildfires, volcanic eruptions, and 33 335 event-free acquisitions to enable thermal hotspot detection and general classification applications. This dataset and associated toolkits provide the community with both an immediately useful resource to speed up future research on energy-efficient preprocessing algorithms and AI-based end-to-end processing systems on board EO satellites.

Index Terms—Onboard artificial intelligence (AI), raw dataset, Sentinel-2, volcanic eruption, wildfire.

I. INTRODUCTION

THE ability of artificial intelligence (AI) to autonomously glean insights from remote sensing data has sparked research into its deployment onboard spacecraft [1]. Furthermore, its application has proven effective in curtailing downlink data rates by preventive filtering of cloud-covered, corrupted data [1], [2], or with higher image compression rates [3], [4].

Most of the previous works relying on AI onboard Earth observation (EO) satellites were demonstrated on high-level satellite optical products [5], [6], [7], [8], which exploit calibrated, orthorectified, and further processed images to lower instrumentation noise and other sources of radiometric distortion. Since such processing chains are designed for ground-based use, they are unsuited for onboard applications. The computational burden of the ground-based processing chain renders its implementation on small-satellites—especially nanosatellites—impractical. Furthermore, it is essential to note that this processing chain alters the raw signal and is susceptible to potential information degradation and loss. Finally, optimizing preprocessing steps is necessary to minimize the computational complexity, potentially reduce the requirement for extra onboard hardware, and enhance mission duty cycles.

In this scenario, a relatively limited number of works have explored the capability of AI models to directly process raw data with minimal preprocessing to reduce the need for computationally intensive onboard operations [9], [10]. This is primarily due to the scarcity of literature containing datasets made of raw data and methodologies to produce such datasets for applications relevant to onboard AI processing. Indeed, increasing the availability of datasets comprising raw imagery could potentially diminish the reality gap prevalent in contemporary studies. This could be achieved by facilitating the training of AI models on satellite data, processed via methodologies that are congruent with the constraints of onboard power and memory capacities. Furthermore, this could stimulate research into more efficient processing techniques. These lightweight techniques could offer enhanced tradeoffs between energy consumption, processing time, and output quality, as compared to solutions that were initially conceived for terrestrial applications.

This work addresses the limited availability of raw imagery for multispectral pushbroom imagery. To this aim, we propose for the very first time a methodology to datasets for the detection of "events" or target objects (e.g., thermal anomalies, vessels) applied to Sentinel-2 raw data and that can be excited to other multispectral EO pushbroom raw imagery. To this goal, the proposed approach capitalizes on existing datasets and algorithmic solutions designed for Level-1 C (L1C) products to extract and identify the corresponding raw images containing the target events and speed-up the dataset creation.

In the case of Sentinel-2 data, "raw" refers to decompressed Sentinel-2 Level-0 (L0) data with additional metadata, which better represents the "raw" data produced by the multispectral sensor.

To automate and make our methodology fully reproducible, we implemented an open-source Python package named *Python for RAW Sentinel-2 data (PyRawS)*[1] that includes utilities for Sentinel-2 raw and L1C data processing.

To showcase the validity of the proposed methodology, we applied our methods for the realization of thermal hotspots in raw Sentinel-2 data (THRawS), a dataset comprising a collection of Sentinel-2 raw data for detecting warm temperature hotspots. We considered this specific application because of its potential to enhance early warning systems and prevent fires from spreading [6], [11], contributing to the safety and security of people in high-risk areas and critical infrastructures, e.g., power lines and oil pipelines.

Since the objective of using raw data is to foster the research of lightweight preprocessing methods, for the implementation of THRawS, we leveraged a lightweight CSC approach. Such coarse spatial coregistration (CSC) method has the potential to be applied on board because of its timing and accuracy tradeoffs.

The rest of this article is organized as follows. Section II provides a state-of-the-art analysis of the processing pipelines designed for multispectral pushbroom imagery, differentiating between missions performing processing on the ground and on board. Moreover, an insight on the availability of raw multispectral data and their use in current missions leveraging AI onboard EO data is provided. Section III overviews the proposed dataset creation methodology. Section IV showcases how we applied the proposed methodology to design THRawS. Section V presents the results in terms of the number of thermal hotspot patches included in the THRawS dataset. Moreover, it provides a description of their spatial and time coverage. Section VI discusses our results and the suitability of using raw data for future usability for onboard AI applications. Finally, Section VII concludes this article. Finally, we provide more details on the Sentinel-2 mission and a comparison in terms of accuracy and latency of the CSC technique adopted to design THRawS to other band-to-band (B2B) alignment solutions in the Appendix.

II. BACKGROUND AND MOTIVATION

A. Multispectral Data Processing Chain

In missions such as Sentinel-2 [12] or Landsat-8 [13], significant postprocessing is applied to enhance satellite multispectral pushbroom imagery prior to their use for the creation of high-end applications-specific products. Typical correction steps include radiometric, geometric, and orthorectification algorithms [14]; most of these steps are applied at ground-segment level. Since our dataset methodology targets Sentinel-2 imagery, this section dedicates particular focus to Sentinel-2 data. However, although the implementation of such processing solutions and product nomenclature differ among various missions, fundamental operations, such as band coregistration, image georeferencing, radiance to top-of-atmosphere (TOA) reflectance conversion are common steps for many of these data processing chains [15][2].

For what concerns Sentinel-2 images are downloaded and processed to be distributed to the end-users through a processing chain, whose schematic representation is depicted in Fig. 1. Data sensed during a given acquisition are organized into granules, each one encompassing the area captured by each of the 12 detectors in a 3.6 s interval [15]. Since the satellite operates in pushbroom mode, a granule encapsulates data from all the bands a sensor acquires. Data produced by the sensor are equalized and compressed onboard before the downlink.

After the download, data available on the ground are defined as L0. The latter are, first, processed to produce metadata, including geographical information, a quick look, and other ancillary information. Then, with their additional metadata, L0 data are decompressed.

In the frame of this article, the data obtained at this point of the processing, including metadata, are defined as "raw data." Level-1 A (L1A) products are then generated by applying a coarse spatial band coregistration process to the raw data. This is followed by processing L1A data to compute radiometrically corrected radiances through various processing steps, including—among others—the inversion of onboard equalization. In addition, the radiometric corrected geometric model is refined and appended to the radiometric corrected L1A data to produce Level-1B (L1B) products.

Lastly, L1C data are derived from L1B products through a process of geometric correction, which includes subpixel multispectral spatial registration, orthorectification, and calculation of TOA reflectances. Cloud and land masks are also generated within this step. Various geometric and radiometric periodic calibration activities are performed to ensure high-quality products, as detailed in [16].

Note that L1C products are delivered for predefined 100×100 km^2 *tiles*, and a given acquisition may have partial or complete data coverage of a certain area of interest (AOI) in a certain raw data granule. Therefore, in general, multiple L1C tiles shall be mosaicked to cover the whole raw AOI.

Since the inversion of the onboard equalization is performed only at L1A data level, its effects are still present for raw data. Similarly, as previously mentioned, the equalized satellite data has undergone on board a process of equalization before compression [12], which aims at minimizing the effect of the wavelet-based compression algorithm on the accuracy/integrity of the detector measurements provided after decompression on the ground. The effects of onboard compression cannot be compensated because of the lossy nature of the used compression scheme. Consequently, eventual distortions due to onboard compression and calibration represent the main difference between raw and data produced by the sensor.

---

[1]PyRaws, GitHub repository. Data is available online at: https://github.com/ESA-PhiLab/PyRawS. Last accessed on 22 June 2024.

[2]Landsat 8 (L8) Data Users Handbook. [Online]. Available at: https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1574_L8_Data_Users_Handbook-v5.0.pdf. Last accessed on 22 June 2024.
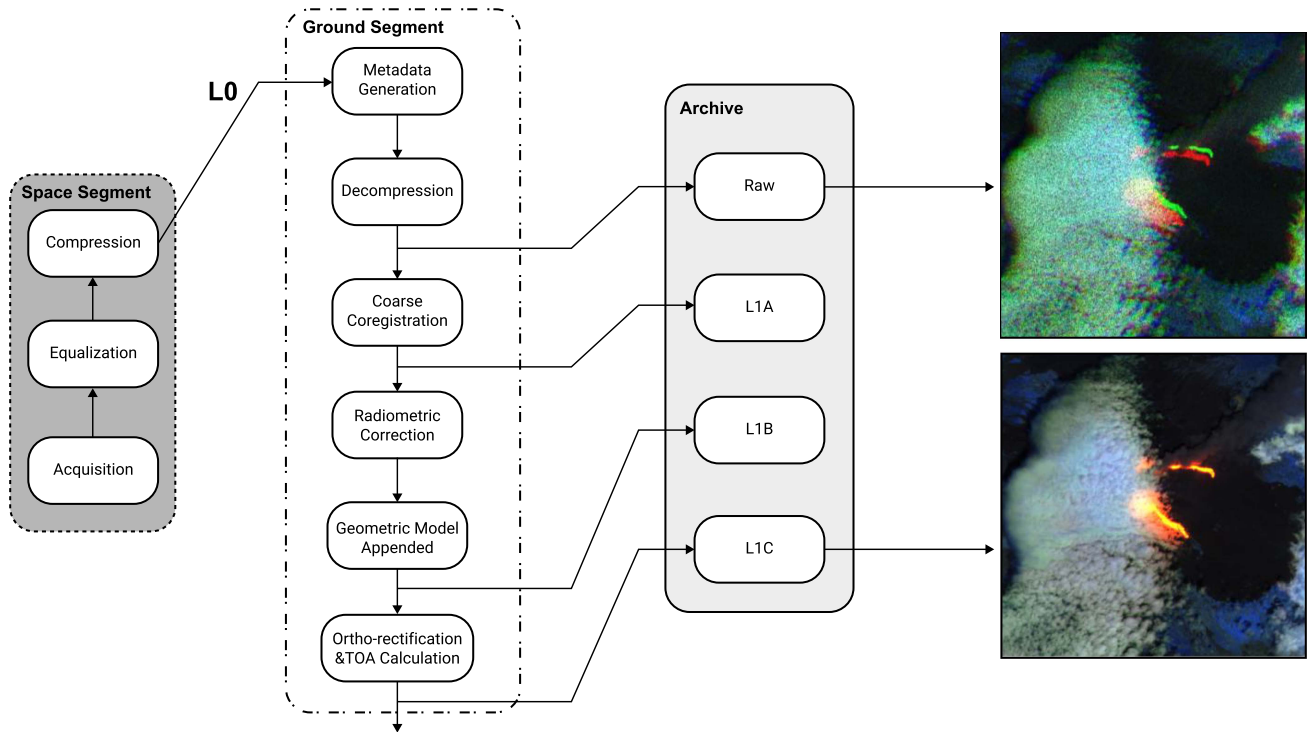
Fig. 1. Illustration of the Sentinel-2 processing chain from satellite data to L1C data. The images in the raw and L1C format show an eruption of the Etna volcano, Italy, for the two processing levels as RGB-like images.

Note that the products in the chain from L0 to L1B included have not been available to users at the time of writing, whereas L1C, atmospheric-corrected bottom-of-atmosphere (BOA) Level-2 A data as well as higher level derived products for different domain applications, such as marine ecosystem monitoring[3] and land cover/land use assessment[4] are available from Copernicus and ESA projects. However, the release of lower-end Sentinel-2 products to the public is foreseen. Products including and beyond Level-2 A (L2A) are out of the scope of this study, which focuses on the combined use of raw and L1C data to create a dataset made of raw data.

B. Multispectral Processing Chains for EO Missions Leveraging Onboard ML

The application of AI on board satellite constitutes an established scientific and commercial venture. In particular, the remote sensing community has heightened the potential to directly handle unprocessed data through the techniques of machine learning (ML) [17] and deep learning (DL) [2], [6]. Prior to the application of ML models, current and upcoming missions predominantly favor the application of preprocessing schemes as a prerequisite [6], [7], [8], [18]. Such solutions usually consist of streamlined workflows that apply only a simplified subset of

geometric and radiometric corrections of the ones performed on the ground.

An illustrative example is the case of the 6-U Φsat-1 [2] Cube-Sat, the first satellite inferring convolutional neural networks (CNN) on a commercial off-the-shelf (COTS) edge device, i.e., the Intel-Movidius Myriad 2 visual processing unit (VPU). Its target application was to perform onboard pixel-level cloud detection by processing three selected bands of the hyperspectral cube. The image processing workflow involves creating the hyperspectral data cube and undertaking B2B spatial registration on three selected bands. Another example is the hyperspectral smallsat for ocean observation (HYPSO-1), a CubeSat equipped with a hyperspectral payload targeting the observation of ocean color to detect algal blooms along Norwegian coasts [17]. The hyperspectral image processing chain on board HYPSO-1 uses the methodology described in [19], which implements radiometric and geometric corrections as linear operations.

A more complex processing chain is planned for the upcoming Φsat-2 mission [20]. Φsat-2 imager is providing seven multispectral bands covering near-infrared (NIR) and visible ranges with 4 m spatial resolution and a panchromatic band (submeter spatial resolution). The payload processing chain processes raw data and provides three processing levels: L1A, consisting of unregistered and no-georeferenced radiance data, L1B, which are fine-georeferenced and fine registered radiance data (<10 m root mean square error), L1C, which are registered and georeferenced TOA reflectance data. Differently from Sentinel-2 data, L1C data are not orthorectified. Given the unavailability of Φsat-2 imagery, emulated sensory data at L1A, L1B, and L1C levels

were provided to the competitors of the *Orbital AI* challenge[5] to investigate different applications for onboard AI. One of the main novelty of Φsat-2 and Orbital AI is the investigation of lower end products for onboard ML applications. However, no emulation of raw data is provided to the end users nor a use of raw data is foreseen in the mission concept of operation.

As stated above, current approaches engage data processing with the intention of reconstructing higher level products onboard, thereby not directly exploiting raw products. The deficiency of raw datasets and the unavailability of methods to treat them inadvertently ease the pathway for higher level products onboard processing. Nonetheless, the adoption of such an approach contributes considerably to the computational cost within the processing workflow and introduces temporal latencies. Indeed, following the empirical assessment of the inference time-lapse on various hardware accelerators, such as NVIDIA Jetson Nano, Google Coral, and Myriad X, Garcia et al. [21] asserted that the time overhead associated with the initial processing stages of Sentinel-2, encompassing B2B alignment—i.e., spatial registration of a band with respect to a reference band—radiometric calibration, and granular coregistration, merit a scrupulous evaluation.

To the best of the authors' knowledge, the first work to directly address the problem of using raw data in an end-to-end fashion for onboard ML applications was conducted in the frame of "the OPS-SAT case" competition [22], [23]. This data-centric competition hosted on the European Space Agency (ESA) Kelvins platform was investigating few-shot learning for onboard satellite applications [24]. The competition was based on the images produced by the OPS-SAT satellite, equipped with an $80\,\text{m} \times 80\,\text{m}$ resolution RGB via on-sensor Bayer pattern camera [25]. In this case, only one frame is acquired from a single snapshot by placing a color filter on top of every sensor pixel. A so-called demosaicing or debayering process is applied to reconstruct a full-color image from a sampled one-color pixel camera. Thereby, images were not affected by misregistration errors due to multispectral B2B alignment processes, obviating one of the most critical preprocessing constraint for the data stack's formation.

Similarly, the work by Fanizza et al. [10] investigated the effect of band misalignment for ship detection on 768×768 pixels RGB images from the Airbus Ship Detection dataset was used for both training and testing. Because of the lack of raw data, the band misaligment was emulated.

To the best of the authors' knowledge, our previous work [9] is the first work investigating the feasibility of using multispectral raw imagery. To this aim, our previous work [9] relies on a dataset containing Sentinel-2 raw coarsely-registered granules to perform onboard vessel detection through an end-to-end pipeline. Despite the preliminary findings of the work show that processing raw data in an end-to-end fashion is possible, the difficulty of extending these results for different applications is raised by the lack of a methodology to automate raw dataset creation and labeling. Indeed, the dataset used for our study was

manually labeled by using a Sentinel-2 L1C dataset as reference. Because of that, the methodology proposed hereby extends our previous work by enabling an automated labeling of raw data by using annotations on the correspondent L1C tiles.

Despite not addressing this problem specifically, the need for raw data availability also emerges from the work [21], which describes the WildRide mission. The latter aims to illustrate that the onboard application of AI is instrumental in minimizing latency, thereby, accelerating the delivery of essential flood mapping information. The DL-based model utilized to perform flood detection had to process RGB data provided by a miniaturized camera.

Nonetheless, given the lack of raw data, Worldfloods was initially trained on Sentinel-2 L1C images by using all 13 bands. As a result, Worldfloods faced challenges arising from the data-shift problem due to the significant difference in terms of radiometric and spectral resolution between the Sentinel-2 sensor and the onboard RGB imager. Therefore, the authors retrained the model on four acquisitions from the used RGB camera to obtain acceptable performance on the raw uncalibrated RGB images, which were manually labeled after the satellite's deployment. This underlines the importance of having raw satellite datasets for the training of AI models to take into account: 1) the impact of the lack of calibration, 2) domain-gap problems, 3) misregistration errors, and 4) other nonideal effects affecting raw data.

III. DATASET CREATION METHODOLOGY

The process of creating a Sentinel-2 raw dataset for the specific purpose of detecting "events" or target objects can be comprehensively understood through the graphical representation depicted in Fig. 2, which showcases all three steps required to create the dataset. The first step is procuring a list of events from online databases or other sources. Each event in the list shall correspond to one or more raw granules. To download such raw data, it is necessary to specify a polygon marking the AOI surrounding the event and a range of dates (start acquisition - end acquisition). Such a polygon used to specify the raw granules is created by visually inspecting the L1C products. This procedure leads to the download of all the compressed raw granules whose reference band (B_{02}) intersects the polygon and whose sensing date is included in that range.

Appropriately selecting the area of the polygon requires considering that for a specific point, the entire collection of the bands could be located in different granules before the spatial registration. Therefore, to collect all the bands for the events of interest, one can use a polygon area of $28 \times k$ km^2 centered at the coordinates of the events, where k is application-specific, which is equal to the double of the maximum bands parallaxes for a single acquisition. For what concerns the parameter k, it shall be higher than the maximum West–East size of the target events. A size of 14 km corresponds to the maximum along-track distance commonly covered by all spectral bands of multispectral instrument (MSI). Therefore, since the satellite track approximately extends from north to south with a smaller closer to the Equator, selecting a polygon that extends 14 km

[5]Orbital Φsat-2. [Online]. Available at: https://platform.ai4eo.eu/orbitalai-phisat-2. Last accessed 22 June 2024.
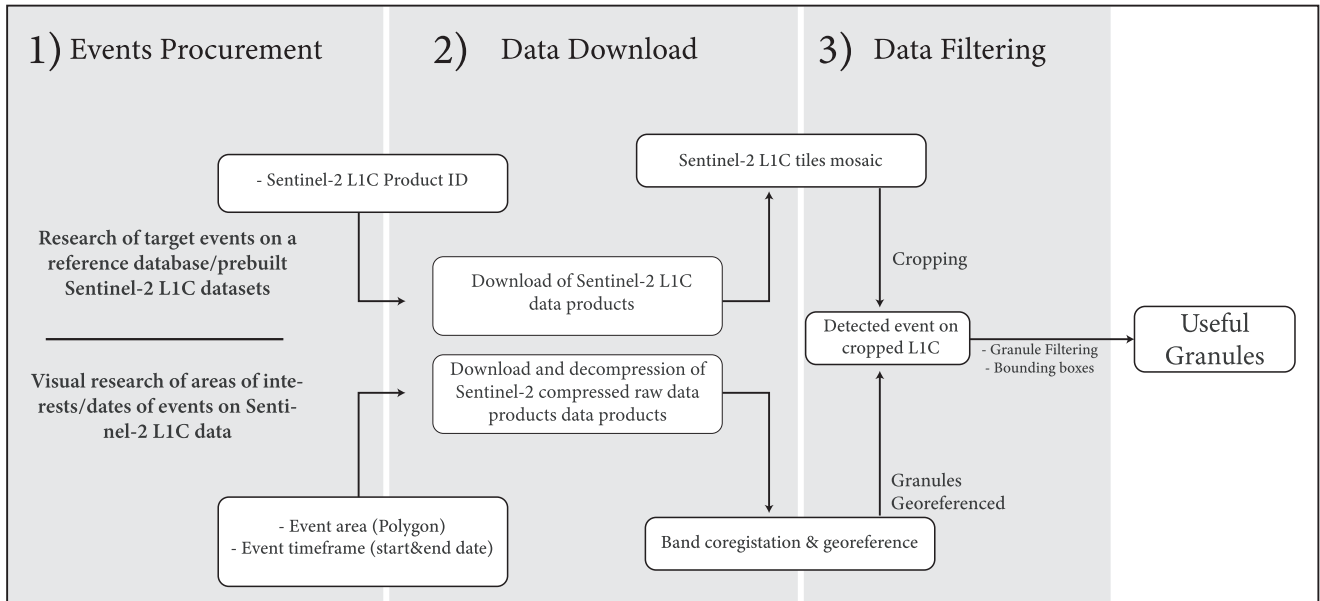
Fig. 2. Overview of the dataset creation methodology consisting of three main steps. 1) Procurement of the list of events by visual inspections of other existing databases, 2) data download, and 3) filtering of useful granules by using state-of-the-art algorithms designed for L1C data, which are mosaicked and cropped on the areas of the raw granules of interest.

vertically toward north and south will ensure that all the bands for the center of the event will be downloaded.

Once the first step is complete, it is possible to download the raw granules and all the related L1C products intersecting the download polygons for each granule. After downloading, the compressed raw data shall be decompressed and reformatted as TIFF files. The download and the decompression of the interest data lead to the conclusion of the second step.

However, the granules downloaded through this procedure generally cover an area higher than AOI. Indeed, since the polygon is not exactly oriented along the satellite track line and because the real AOI could be much smaller than the polygon, this procedure generally leads to a higher number of raw data granules than required. This can be seen in the left image of Fig. 3, which shows the granules downloaded (rectangular shapes numbered 0 to 5) for the volcanic eruption of Etna on 30 August, 2021 included in the THRawS dataset. The polygon used to specify the AOI is shown in light blue. But, clearly, only the raw data granules in green and red actually contain the event. Therefore, there is the need to identify and label the granules containing the events of interest.

To this aim, we conceived a methodology that filters the raw data granules containing the events of interest by processing L1C tiles by relying on the availability of state-of-the-art algorithms for L1C products.

Such state-of-the-art algorithms usually exploit a subset of the 13 bands included in Sentinel-2 data. Let us define as $B_S \triangleq [B_x, B_y, \ldots, B_z]$ the ordered list of the bands used by the identified algorithm, where B_x is the first band of the collection. Then, for each event in the dataset, every granule is processed as follows. First, we perform the coregistration of the different bands of B_S with respect to B_x. Spatial coregistration algorithms are designed to eliminate displacement between

bands, errors in skewness, rotation, and warping [26]. In our case, this step is mainly necessary since the bands in a granule feature a displacement due to the pushbroom nature of the sensor and additional nonideal effects [15]. Such band shifts will make it difficult to retrieve and visually inspect events on raw data after being identified. By applying the band spatial coregistration with respect to B_x, among other effects, all the bands of B_S will be shifted along and across track to match the area covered by the band B_x. The missing pixels due to the shift procedure could be cropped, filled by zeros, or by pixels of adjacent granules when available. After applying the band coregistration, we performed georeferencing of the first band in the granule. This step is needed since only the coordinates of the corners of the entire granule footprint are provided in the granule metadata, but not for the different bands [15].

Once the band B_x has been georeferenced, it is possible to crop L1C data on the same area of its band B_x. To this aim, all the L1C tiles intersecting the polygon used to download raw data are mosaicked and cropped on the same area of the band B_x. To this aim, the information provided by B_x coordinates as a common reference for L1C and raw data.

After cropping, these adjusted L1C tiles are processed using the identified reference algorithm to search for events of interest. One or more bounding boxes are created on the L1C data for each event. When finding a bounding box is impossible, the correspondent granule is discarded. Otherwise, the bounding boxes are warped back on the granule using an affine transformation and manual fine-tuning.

We mark *useful granule* as a granule whose band B_x includes at least one bounding box. It is necessary to remark that this definition of *useful granule* depends on the band collection used. In the case of the volcanic eruption "Etna_00" of the THRawS dataset shown in Fig. 3, the volcanic event is included
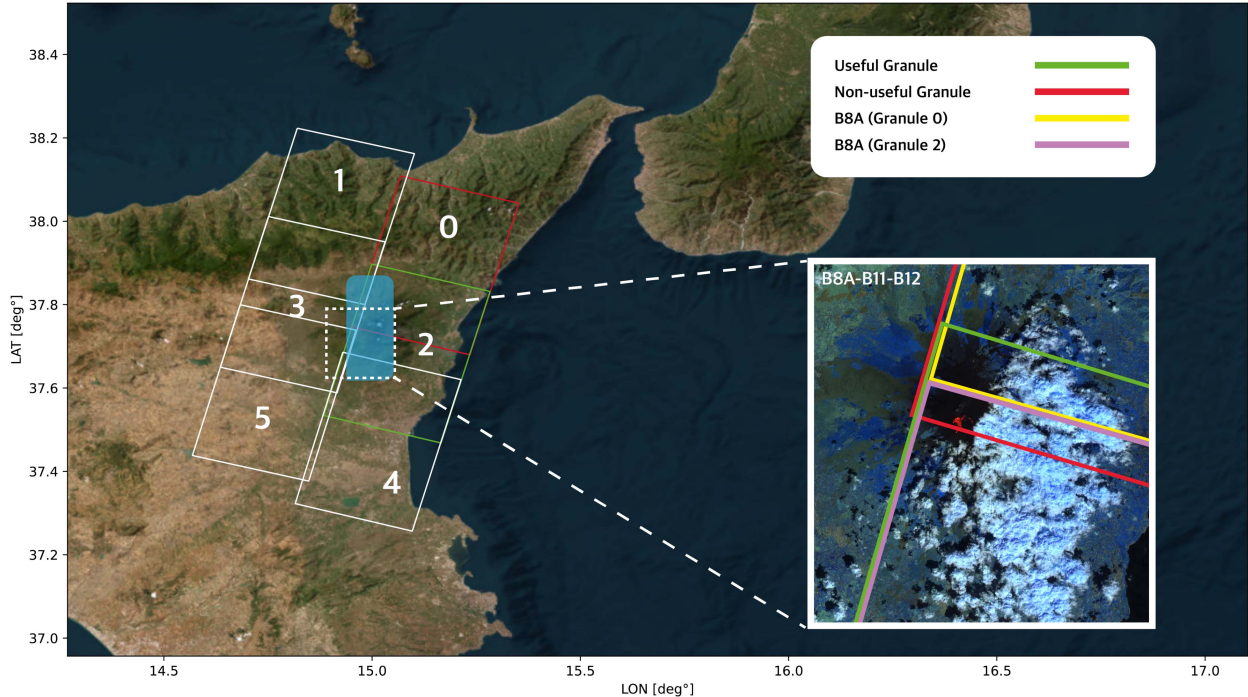
Fig. 3. Example showcasing the downloaded granules for a specific event of the THRawS dataset. The polygon used for data retrieval (light blue) intercepts several granules marked in red, green, and white. The volcanic eruption is included only on the granules whose boundaries are marked in green and red. On the right, a zoomed view shows the band B_{8A} (the first of the collection) of the green and red granules in yellow and pink. Since the band B_{8A} of the green granule only includes the volcanic eruptions, the green granule is the only "useful granule." On the contrary, despite the red polygon partially surrounds the event of interest, this happens for other bands than B_{8A}. Because of that, the event is "nonuseful" for our definition of the band collection $B_S = [B_{8A}, B_{11}, B_{12}]$.

in both the granules 0 and 2, respectively, marked in red and green. However, since for the THRawS dataset $B_x = B_{8A}$, the procedure previously described will lead to selecting only granule 2 as a *useful granule*. Indeed, when considering the area covered by the band B_{8A} in granule 0 (yellow rectangle), one can see that it does not include the volcanic eruption[6].

IV. THRawS Case Study

This section demonstrates how we applied the proposed methodology described in Section III to create the THRawS.

As mentioned in Section III, the first step to creating a dataset according to the proposed methodology is to identify the event's location and time ranges by relying on existing datasets or online databases. More specifically, THRawS contains volcanic eruptions and wildfire events, for which we identified two different sources.

The starting point for selecting tiles with eruptions is the Smithsonian Institution database [27], providing detailed information on global volcanism. From the database, we selected only events occurring during the Sentinel-2 A (S2-A) launch until the time of writing, i.e., from 2016 to 2022, and we aimed to sample different latitudes and seasons. From the database, we only selected more intense and explosive eruptions, with volcanic explosive index (VEI) larger than 1, since effusive

eruptions are less dangerous [28]. As events can last for several years, we initially selected up to three volcanic events for each eruption and up to three eruptions for a single location to ensure the diversity of the dataset. However, during the revision of the dataset, some raw data granules originally marked as Not-events were found to contain volcanic thermal hotspots, leading to more than three volcanic events for some of the eruptions.

For shorter (days-months) eruptions, we selected acquisitions as close as possible to the event starting date to test the detection capability in the early stages. From the list of eruptions, we visually inspected the candidate L1C tiles to confirm that the event was captured. For each of the selected L1C, we marked the location of the volcanic event. We used its coordinates as a center of the rectangular polygon to download raw data granules, as detailed in Section III. In particular, we selected the horizontal size of the download polygon to be $k = 10$ km since it was sufficient to contain the broadest warm temperature hotspot in the dataset.

Compared to volcanic eruptions, fire events are challenging to detect as they are short-lived and their locations are not known in advance compared to volcanic eruptions. We compiled a list of events starting from the Copernicus Emergency Monitoring System database[7] and integrated the resulting list of events with real-time detection services[8] and information available online

---

[6]In the THRawS dataset, granules 4 and 5 of the event *Etna_00* are also marked as *useful granules* since they contain small warm thermal hotspots that are not visible in the image. This fact is omitted in the text for the sake of explanation.

[7][Online]. Available at: https://emergency.copernicus.eu/mapping/list-of-activations-rapid; last accessed on 22 June 2024.

[8][Online]. Available at: https://firms.modaps.eosdis.nasa.gov/map; last accessed on 22 June 2024.

from space and environmental agencies. As for eruptions, candidate L1C acquisitions have been visually inspected to verify the presence of a firefront or extensive smoke. Finally, we retained a number of fire events roughly equal to that of eruptions to balance the dataset distribution.

Both wildfire and volcanic eruptions events were selected by visual inspection independently and, then, crosschecked by two experts to ensure reliability and mitigate the effect of human biases. In addition to granules containing L1C data, we specifically introduced hard negatives to optimize model performances by a proper selection of Not-events. This approach aims to establish consistent relationships between thermal anomalies and network outputs as it enables the identification of specific anomalies in the image not necessarily related to volcanic eruptions, e.g., crater rim edges. Selecting Not-events involves visually inspecting each granule and manually selecting to ensure the absence of hotspots. It is worth noting how this is important to allow for a more accurate ingestion of the data in deep neural networks (DNN) models.

After selecting the reference warm temperature hotspots events, we proceeded with the download and filtering of the corresponding raw and L1C data.

In particular, we selected $B_S = [B_{8A}, B_{11}, B_{12}]$, given the choice of the reference algorithm to detect warm thermal hotspots in thermal anomalies by Massimetti et al. [29].

The next required step is the coregistration of the different multispectral bands in the bands collection B_S. Indeed, given two generic bands B_n and B_m of a granule, the band B_n is generally shifted of a certain number of pixels $S_{B_n\_B_m}$ with respect to B_m. In general, it is possible to assume that $S_{B_n\_B_m}$ is made of two components as shown in the following:

$$S_{B_n\_B_m} = \overline{S_{B_n\_B_m}} + \Delta S_{B_n\_B_m} \qquad (1)$$

where $\overline{S_{B_n\_B_m}}$ is the systematic shift due to pushbroom acquisition mode and additional offset [15], whilst $\Delta S_{B_n\_B_m}$ is an aleatory component due to mechanical vibrations and other nonideal effects.

Spatial coregistration algorithms are generally designed to eliminate both the components of $S_{B_n\_B_m}$ so that the spatial displacement between B_n and B_m can be minimized [30], [31], in addition to errors in skewness, rotation, and warping [26]. To achieve these objectives, as detailed in Section II, spatial coregistration techniques rely on a suite of algorithms that are generally computationally intensive and could require dedicated additional hardware to be implemented on board satellites [2], [30], [31], [32]. Since one of the objectives of this work is to foster the study of lightweight and energy-efficient preprocessing algorithms, to design the THRawS, we adopted a coregistration solution designed to compensate only the systematic shift error component $\overline{S_{B_n\_B_m}}$ by shifting B_m of a fixed value equal to $\overline{S_{B_n\_B_m}}$.

To this aim, as for [19], our solution uses one-time precalculated spatial shift values to compensate for the average spatial displacements both along and across the satellite track due to the sensor's pushbroom nature and additional sensor offsets. An example of this precomputed calculation is reported in Table IV in the Appendix. The thus developed approach does
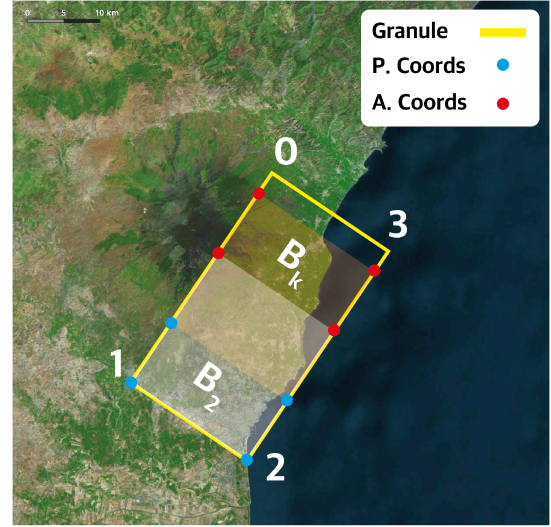

Fig. 4. Pictorial view of a granule with Prior and Afterwards coordinates.

not require using ground control points nor other ancillary data. In addition, it does not require to perform keypoints extraction and matching, typical of features-based technique [33], [34], [35], [36], [37]. These shift values do not consider effects due to noise, attitude disturbances, or depending on the surface characteristics. However, as shown later in Table IV, for Sentinel-2 data $\overline{S_{B_n\_B_m}} \gg \Delta S_{B_n\_B_m}$, especially for the along-track component (*along-track shift*). Owing to its limited precision, we term our coregistration technique "coarse." As outlined in Appendix, the average registration errors remain smaller than the smallest identified thermal hotspots. Because of that, the use of the proposed CSC does not lead to loss of events. In addition, the proposed approach is lightweight, which makes it promising for onboard satellite applications and accelerates the dataset creation. The methodology to estimate the shift values used for the CSC is detailed in Appendix.

After applying the band coregistration, we adopted a custom band coarse georeferencing (CG) scheme based on the fixed B2B displacement values used for the CSC applied to the granule corner coordinates to determine the coordinates of each band's corners. Being a lightweight solution, this approach is promising to be adopted for onboard satellite applications. However, it assumes that the coordinates of the four granule corners are known. This information is contained in the granule metadata and requires additional processing steps to be made available on board, which were not investigated in this study.

Although we applied the following methodology to band B_{8A} only, our approach is general, and, therefore, we provide a description for a generic band $B_k \in B_S$. We initiate this process by extracting the coordinates of the four corners of a granule, which constitute the essential ancillary information provided in the metadata of the Sentinel-2 raw product. Referring to Fig. 4, we define two key variables for our methodology. The first, prior coordinates (PC), represents the initial two corner coordinates scanned by the pushbroom system. The second, afterward coordinates (AC), denotes the final two corner coordinates scanned.

Given the known location of PC for band B02 (PC_{B02}), our method utilizes the coarsely estimated offsets $S_{B_k\_B_{02}}$ from this reference band to determine the PC, PC_{B_k} for any arbitrary band B_k.

For a band B_k, we can estimate the offsets from the reference band using the following:

$$PC_{B_k} = PC_{B02} + |S_{B_k\_B_{02}}| \cdot \frac{R_{B_k}}{R_{B_{02}}} \cdot \frac{G_a}{G_L} \qquad (2)$$

where $S_{B_k\_B_{02}}$ represents the shift value between band B_k and the reference band $B02$, G_a is the granule length expressed in arc of latitude/longitude, prompted from the polygon coordinates at positions 0 and 1 of the granule (see Fig. 4), G_L is the granule length expressed in px at 10 m resolution. Essential to note is that, given the variations in spatial resolution between bands, the shift values are scaled specifically for each band with the resolution scale factor $R_{B_k}/R_{B_{02}}$, where R_{B_k} and $R_{B_{02}}$ are the spatial resolutions of band B_k and $B02$, respectively. This equation allows us to calculate the PC for each band, facilitating the process of band georeferencing.

The AC are, instead, prompted using ancillary information by applying the following linear proportion:

$$G_L : B_{l_k} \cdot \frac{R_{B_k}}{R_{B_{02}}} = G_a : \Delta_{B_k} \qquad (3)$$

where B_{l_k} is the length of the k band expressed in px.

Solving the proportion in (3) gives the offset Δ_{B_k} referred to a generic k band as follows:

$$\Delta_{B_k} = \frac{R_{B_k}}{R_{B_{02}}} \cdot \frac{B_{l_k} \cdot G_a}{G_l}. \qquad (4)$$

In the end, the AC can be calculated as follows:

$$AC_{B_k} = PC_{B_k} + \Delta_{B_k}. \qquad (5)$$

The set of AC and PC enables the coarse georeferencing of each pixel within the k spectral band.

After computing the coordinates of the band $B_k = B_{8A}$, it is now possible to crop collate the various L1C products and crop them on the area of the band B_{8A}. To retrieve the same area on the L1C products, for each band in B_S, we extract and collate the correspondent band in each of the L1C tiles that intersect the polygon used to download granules. The resulting mosaic is then cropped on the AOI by using the geographical coordinates retrieved by using the CG scheme. Then, the cropped mosaic of the various bands is resampled to match the band with the coarsest resolution.

Finally, to detect thermal hotspots in L1C data, we used a simplified implementation of [29] to generate a hotmap containing thermal hotspots and extract bounding boxes. This algorithm relies on the bands $B_S = [B_{8A}, B_{11}, B_{12}]$, which we selected as band collection. Therefore, we cropped and mosaicked the bands in B_S of the L1C products using the coordinates of the georeferenced and coarsely registered raw granules. To spot thermal hotspots on L1C data, as for the original implementation of Massimetti et al. [29], we derived a hotmap containing candidate anomalies fulfilling four logical conditions that are calculated by applying fixed thresholds on the ratios of TOA reflectance of the different bands.

More specifically, as in the original work [29], the logical equation to activate a pixel p in the hotmap is as follows:

$$p = \alpha \oplus \beta \oplus S \oplus \gamma \qquad (6)$$

where $\oplus$ is the logical sum between the logical conditions α, β, S, and γ. The latter are calculated as a function of the TOA reflectance values of the bands B_{8A}, B_{11}, and B_{12} (ρ_{8a}, ρ_{11}, and ρ_{12}) as follows:

$$\begin{cases} \alpha = \left(\frac{\rho_{12}}{\rho_{11}} \geq 1.4\right) \otimes \left(\frac{\rho_{12}}{\rho_{8A}} \geq 1.2\right) \otimes (\rho_{12} \geq 0.15) \\ \beta = \left(\frac{\rho_{11}}{\rho_{8A}} \geq 2\right) \otimes (\rho_{11} \geq 0.5) \otimes (\rho_{12} \geq 0.5) \\ SR \triangleq \text{SUR}\{\alpha \oplus \beta\} \\ S = [(\rho_{12} \geq 1.2) \otimes (\rho_{8A} \leq 1)] \\ \qquad \oplus [(\rho_{11} \geq 1.5) \otimes (\rho_{8A} \geq 1)] \\ \gamma = (\rho_{12} \geq 1) \otimes (\rho_{11} \geq 1) \otimes (\rho_{8A} \geq 0.5) \otimes SR \end{cases} \qquad (7)$$

where $\text{SUR}\{\}$, $\oplus$ and $\otimes$ are, respectively, the surrounding, the logical OR, and the logical AND operations.

Then, we extracted a bounding box for every cluster of pixels in the hotmap using *scikit-image* [38]. Finally, to minimize the false detection rate, we filtered those bounding boxes surrounding a number of active pixels in the hotmap lower than 9 to discard small clusters of pixels as suggested in [29].

In the original implementation, the authors extracted statistical information related to each cluster of pixels to set cluster-dependent thresholds to reduce the false positive rate. We omitted this step in our implementation since we are not interested in pixel-level information. Indeed, the creation of a hotmap is used only for binary classification and bounding box extraction.

The work by Massimetti et al. [29] was initially implemented for a volcanic eruption. However, other works such as [39] consider volcanic eruptions and wildfires as thermal anomalies using the same Sentinel-2 bands and threshold-based detection methods.

Given the methodological similarity, we also used the approach of [29] to detect fire events in L1C tiles.

After extracting the bounding boxes on the mosaicked and cropped L1C tiles, we reprojected them on the raw products using an affine transformation and manual tuning. Indeed, it is essential to note that the bounding boxes extracted from the L1C products cannot be straightforwardly used for the raw data due to their different projections. While the raw products are represented in azimuth-range coordinates ($EPSG : 4326$), L1C products are in latitude-longitude projected in the *WGS84* system. A further issue is added by the geometrical correction applied to the L1C images, in which a resampling using a 90 m DEM (PlanetDEM 90)[9] is employed. In order to exploit the L1C bounding boxes, we warped the points from the L1C to the raw coordinates with an affine transformation. We defined the transformation matrix from L1C to raw coordinates using the

---

[9][Online]. Available. Reference Document: GMES-GSEG-EOPG-TN-09-0029 (v2.3)
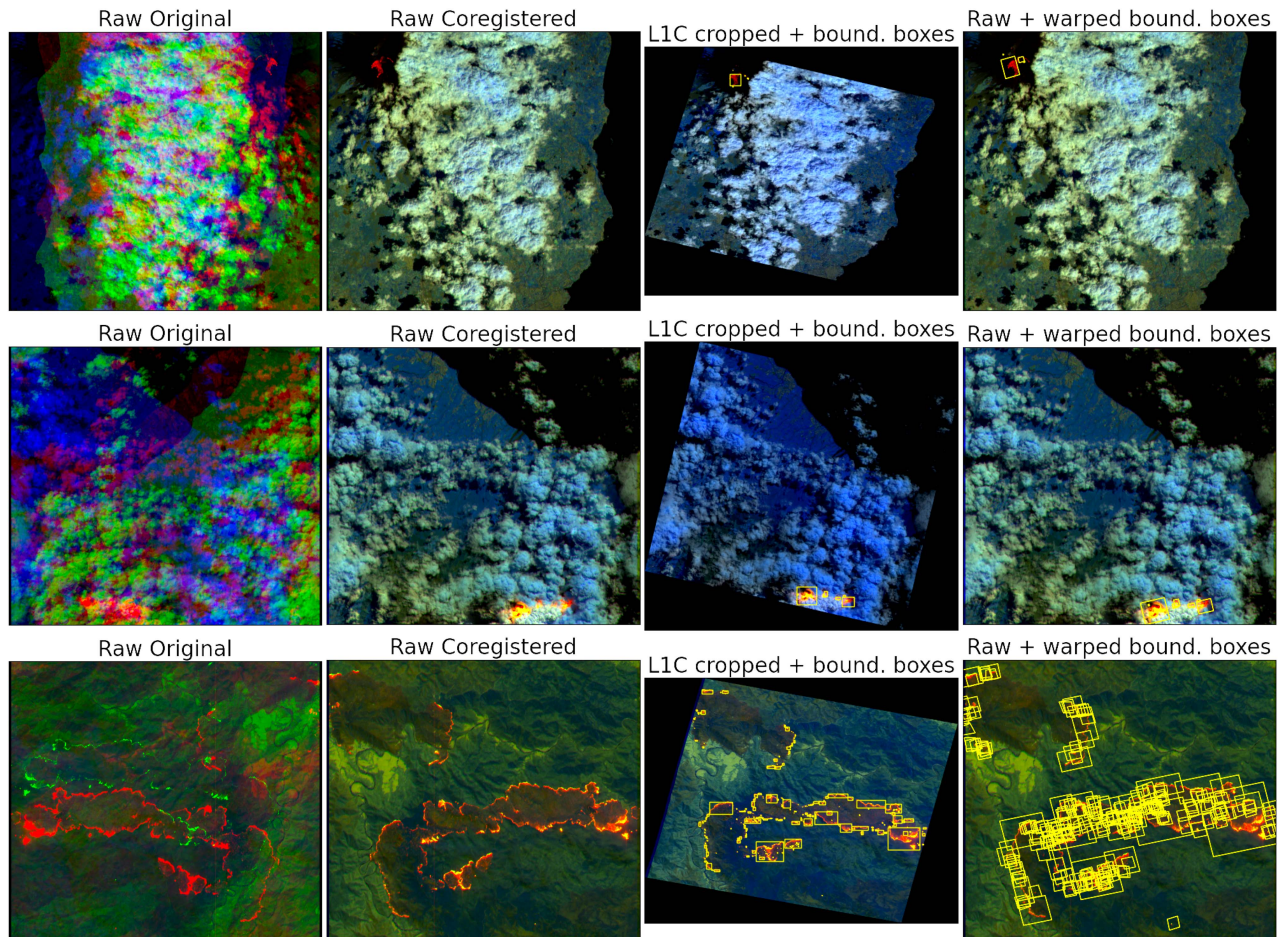
Fig. 5. Each row shows all the processing steps for the bands B_{8A}, B_{11}, and B_{12} of the various raw data granules: First image from left of each row displays the raw data granule, the second image shows a spatially registered granule, the third image shows the correspondent L1C tiles cropped on the coordinates of first band of the raw data granule band collection with the detected bounding boxes, and the rightmost image showcases the bounding boxes warped on the raw data granules.

coordinates of three corner points. After applying the correction, our results showed that there still exists a discrepancy between the bounding box and the actual event in several cases. In order to rectify the issue, we manually tuned the bounding boxes for eruptions and created a buffer for fire events. Finally, we performed a visual inspection to validate each presence/absence of events in the *useful granules*. Except for this step, all preceding raw data processing procedures in our methodology can be automated through the *PyRawS* package, which offers suitable API functions to implement the described CSC, CG techniques, L1C manipulation, and detected events transfer between L1C and Raw data. It also includes the simplified implementation of the Massimetti [29] algorithm previously reported.

An example of applications of the different methodology steps by using *PyRawS* on raw granules and L1C data of THRawS are depicted in Fig. 5.

V. RESULTS

This section provides a quantitative analysis of the THRawS dataset, covering aspects spanning from the composition of its granules to its representation in time and space.

TABLE I
NUMBER OF EVENTS, USEFUL GRANULES, AND DISCARDED GRANULES IN THE DATASET

Product/event	Eruptions	Fire	Not-event
raw products	58	20	10
raw useful granules		135	11
Total raw useful granules		146	
Discarded granules		707	
Total number of raw granules		905	

A. Granule Distribution

As detailed in Table I, THRawS comprises 88 thermal hotspot events initially selected as 58 volcanic eruptions, 20 fire events, and 10 Not-events. It is crucial to notice that each of the 88 elements corresponds to one or more *useful granules*, i.e., granules containing thermal hotspots that were appropriately annotated through the proposed approach. Furthermore, eruption events do not necessarily correspond to granules containing volcanic eruptions only but also wildfires. This is due to using a standard algorithm by Massimetti et al. [29] to identify *useful granules*

Fig. 6. Decomposition of a raw granule into 256×256 patches with 25% of patch overlap.

TABLE II
PATCH STUDY ANALYSIS HIGHLIGHTING THE NUMBER OF EVENTS AND NOT-EVENTS, AS WELL AS THE PROPORTION OF EVENTS, AS THE OUTPUT OF THE DIFFERENT PATCH DECOMPOSITION PROCESS

Overlap (%)	Patch Size	No. of Thermal hotspot samples	No. of Not-Thermal-hotspots samples	Proportion of Events
0.25	**(128, 128)**	**1090**	**33335**	**0.031663**
0.25	(256, 256)	581	6916	0.077498
0.25	(384, 384)	527	3298	0.137778
0.25	(512, 512)	541	1907	0.220997
0.33	(128, 128)	1211	37957	0.030918
0.33	(256, 256)	783	9009	0.079963
0.33	(384, 384)	527	3298	0.137778
0.33	(512, 512)	541	1907	0.220997
0.50	(128, 128)	2158	65315	0.031983
0.50	(256, 256)	1220	14080	0.079739
0.50	(384, 384)	1102	6395	0.146992
0.50	(512, 512)	852	2973	0.222745
0.75	(128, 128)	8345	248848	0.032446
0.75	(256, 256)	4582	50651	0.082958
0.75	(384, 384)	3326	18706	0.150962
0.75	(512, 512)	2189	7603	0.223550

The values used for the generation of THRawS dataset are marked in bold.

Fig. 7. Pictorial view of the geographical distribution of the "fire" and "eruption" raw useful granules in the THRawS dataset.

that do not distinguish between volcanic eruptions and wildfires. This issue could be easily solved by manually reviewing the *useful granules*. However, since the goal of THRawS is to create a thermal hotspot dataset and not to differentiate between volcanic or wildfires, this step was not performed.

Hence, the total number of *useful granules* in the THRawS dataset is 146, of which 135 correspond to thermal anomalies, and 11 granules depict the volcanic areas where no thermal hotspot was observed. Furthermore, when the discarded granules by the proposed methodology are considered, the total number of granules included in the THRawS dataset is 905.

Notably, the number of images containing the events can be significantly increased by cropping the bands $B_S = [B_{8A}, B_{11}, B_{12}]$, whose size is 1152px $\times$ 1296px, into patches of smaller size. This step is typically implemented in current smallsats missions to isolate patches containing different spatial features and ensure meeting the onboard memory requirements by limiting the size of patches provided as input to ML models [4], [2], [22], [23]. Fig. 6 illustrates this concept by demonstrating the effect of employing 256px $\times$ 256px patches with a 25% overlap in raw granule patches decomposition.

Therefore, the number of dataset samples can be significantly increased depending on the dimensions of the patches and their degree of overlap. Table II presents a comprehensive analysis of the impact of patch sizes on the THRawS dataset using varying patch sizes and degrees of overlap.

In the computational methodology employed for this study, a patch is deemed to encompass an event if the spatial extent delineated by its annotation exceeds a threshold of 5 pixels within the patch coverage. As shown by Table II, decomposing the single granules into patches as requested by the typical onboard hardware requirements leads to a significant increment in the number of samples containing thermal hotspots from hundreds (i.e., number of granules) to several thousand (i.e.,

number of patches) depending on the overlap and patch size. To generate THRawS, we opted for 25% overlap and patch sizes of (128 $\times$ 128), leading to more than 1000 hotspot sample patches without significantly increasing the correlation among patches due to large overlap sizes.

However, applying such a patch decomposition strategy would lead to a high imbalance in training patches between the dataset's thermal hotspot/no-thermal hotspot classes. Because of that, class imbalance shall be adequately handled by upsampling the event or downsampling the Non-event class.

B. Temporal and Geographical Coverage

Fig. 7 shows the geographical distribution of volcanic eruptions and wildfires. Volcanic eruptions are primarily concentrated in Central/South America, Africa, Indonesia, the Philippines, and some European islands (such as Sicily, the Canary Islands, and Iceland).

Wildfires scenes are, instead, concentrated in Europe, Australia, Africa, Greenland, and Central/South America. When considering both wildfires and volcanic eruptions, the thermal anomalies contained in THRawS have almost global coverage in the various continents except for continental Asia, North America, and Antarctica.

It is worth noticing that given the nature of thermal hotspots included in the dataset, most of the images focus on rural areas far from urban centers. In particular, wildfire hazards are more likely in areas where abundant biomass build-up in the wet season can be converted into fire fuel during the local dry season [28], such as in tropical countries. Because of that, eventual AI models aiming to detect thermal anomalies in urban areas might require additional complementary granules depicting wildfires in such areas.

For what concerns time representativeness, volcanic eruptions were selected by picking up to three volcanic events for each eruption as defined in the Smithsonian Institution database [27]. In particular, by picking up the Sentinel-2 granules whose sensing times match the day of the eruption starting date, we included volcanic events in their early phase to enable studies on the early detection of volcanic eruption. Furthermore, given the presence of other events whose sensing times are successive months or years to the first event, THRawS features volcanic events in different stages to enlarge its time representativeness. Instead, wildfires in vegetated areas are more likely during the local afternoon since air has warmed and the ground drier than in the morning hours [40]. Given the Sentinel-2 orbit characteristics, this implies that the database will mostly contain more severe and more protracted fire events. In general, the sensing time of events included in THRawS ranges from 16 January 2017 until 01 June 2022, which spans the early stage of the Sentinel-2 sensor life until the dataset creation time, encompassing possible differences in the Sentinel-2 sensors calibration or mission set-up.

C. Data–Label Representativeness

We implemented a comprehensive two-step approach to validate the dataset. In the first step, we conducted a preliminary selection process, followed by a thorough visual inspection of the L1C data. This step was crucial for ensuring the initial quality and relevance of the data. In the second step, we focused on a detailed visual analysis of the *useful granules*. These granules were specifically chosen by a well-regarded reference algorithm, as detailed in the study by Massimetti et al. [29]. This allowed us to assess the practical utility of the data as defined by established algorithmic criteria. Visual inspection was performed by two independent experts. No significant errors in labeling warm thermal hotspots arouse from the application of our implementation of the reference algorithm [29] on L1C products. Any errors in bounding box warping from L1C to raw *useful granules* was manually corrected at the best of the authors' possibilities.

The overarching goal of this two-step validation process was to confirm the accuracy and relevance of the events recorded within the AOI. By cross-referencing with existing datasets, we aimed to reduce the likelihood of false positives, which could arise from the inherent limitations or biases of the reference algorithm. This methodical approach enhances the reliability and applicability of our dataset for further analysis and application.

VI. DISCUSSION AND FUTURE PERSPECTIVES

This section discusses the use of raw multispectral imagery for ML-based onboard applications. Moreover, it discusses the suitability of the proposed methodology to generate datasets with sufficient fidelity in representing sensory-produced data.

A. Perspectives on Usability Raw Data

As previously mentioned, providing raw imagery is to investigate lightweight onboard preprocessing algorithms or alternative end-to-end solutions to reduce latency and energy consumption, which are of utmost importance for early section systems. The idea behind the use of raw data is that ML solutions could be trained to handle possible disturbances due to the use of raw imagery, curtailing the need for high-quality data. Some preliminary studies have already demonstrated the potential of DL models to appropriately handle registration errors [9], [10] by including emulated misregistered data during the training.

In particular, our previous work [9] provides a preliminary demonstration of the feasibility of using coarsely registered Sentinel-2 real raw imagery for onboard vessel detection with coarsely registered images. For this reason, these preliminary results, when combined with the advantageous quality/latency tradeoffs of the proposed CSC technique used to design THRawS (refer to Appendix), represent a valuable example on the value of raw data can be used to investigate lightweight energy-efficient tradeoffs. A detailed benchmarking of different registration techniques in terms of timing and accuracy on THRawS is provided in the Appendix.

In addition to lightweight processing experimentation, similar to our previous work [9], the provided THRawS dataset can be a useful tool for the training ML models for end-to-end processing onboard thermal hotspot detection and classification on Sentinel-2 raw imagery on rural areas.

However, differently from THRawS, the dataset used in our previous study [9] was built via manual labeling because of the lack of an automated procedure to handle dataset creation. In view of that, due to the possible upcoming release of further raw Sentinel-2 data, we believe that the presented methodology could play a fundamental role in fostering future research on raw multispectral imagery by significantly curtailing the time for dataset preparation, whose impact on research in terms of development time is already discussed in numerous studies [25], [41]. As previously mentioned, the fundamental steps of the proposed methodology are made reproducible through our dedicated open-source software *PyRawS*, which facilitates the handling of raw data significantly.

Hence, the provided methodology could be a useful instrument to start filling the gap in raw data availability by providing an automated procedure for dataset creation and labeling on Sentinel-2 imagery and beyond. Indeed, although the present methodology targets Sentinel-2 raw data, its usability is not limited to Sentinel-2 onboard applications, being in principle

applicable to other multispectral pushbroom imagery showing similar processing chains [2], [20], although with opportune modifications.

B. Fidelity in Representing Sentinel-2 Raw Sensory Data

As mentioned in Section II, the proposed raw data are decompressed L0 data with additional metadata. Because of that, they differ from the data produced by the imager in terms of the presence of such additional metadata and the noncompensated effects due to onboard compression and equalization.

We used additional metadata only in the phase of dataset preparation (e.g., the coordinates of the granule's corners). However, we do not require metadata information in the granule processing chain, for instance, to perform the proposed CSC. Regarding the near-lossless compression applied on board Sentinel-2 satellites before the data download, the scheme introduces information losses that affect the other products in the processing chain, such as L1C data. Since numerous works relying on ML for the detection of specific events on L1C data exist [5], [6], [8], it is possible to conclude this information loss shall not hinder the usability of ML on raw data.

From what concerns onboard equalization, it consists of a bilinear function [42] that is applied prior to the wavelet compression to minimize its effects. Assessing the impact of onboard equalization on the performance of a generic AI model and comparing Sentinel-2 raw to sensory data without proper experimentation is harsh. In general, however, being natively designed to be performed on board satellites, such algorithms lead to a low computational burden. Because of that, it is reasonable to assume that such algorithms could be part of the onboard preprocessing pipeline in case their presence will be proven to be fundamental to ensure sufficient performance of onboard ML models.

C. Limitations of the Proposed Dataset Creation Methodology

Manual fine-tuning and visual inspection are still required to correctly transfer the bounding boxes from L1C to raw products. This fact currently represents the main obstacle to fully automate the dataset creation procedure and is due to local deformations in the spatial mesh of L1C products due to orthorectification and geometric calibration procedures. To solve this issue, possible solutions require the matching of local spatial features in the areas of the bounding boxes through key-points identification and matching algorithms, such as SuperGlue [43], LightGlue [44], and others. However, a detailed investigation of this problem is beyond the scope of this work and will be investigated in a future study.

D. THRawS Dataset Limitations

In our study, we identified several limitations that warrant further investigation. The first pertains to the application of the algorithm by Massimetti et al. [29] for detecting thermal hotspots in both volcanic and wildfire events by using uniform threshold value across all events and locations [39]. However, our study did not implement the cluster-based thresholds that were integral to the original algorithm by Massimetti et al. [29]

to reduce the false positive rates. Our choice is motivated by the fact that we do not aim to provide an accurate pixel-level hotmap, but a coarse detection of the events. However, this approach, when combined with our simplified clustering solution and the buffer created on fire events to mitigate the inaccuracies of L1C/Raw warping, could influence the precision of bounding box placements, necessitating additional tuning for applications that demand exact positioning.

A notable shortcoming in our study stems from the uneven representation of wildfire and volcano events within our dataset. While we have included both wildfires and volcanic eruptions to broaden the spectrum of thermal anomalies, our objective does not encompass the differentiation between these types of thermal anomalies. Despite that, we acknowledge that the inclusion of additional wildfire events in North America, Asia, and others with currently limited coverage could significantly benefit our dataset by further enlarging the variety of thermal anomalies included. Moreover, as detailed in Section V-B, the dataset predominantly features severe and prolonged fire events due to the operational characteristics of Sentinel-2, which may limit the representation of less intense fires commonly found in diverse ecosystems such as savannas and forests. Future studies could consider integrating raw data from additional available sensing platforms to enhance the diversity and representativeness of thermal anomaly events in urban and natural environments.

A dedicated follow-up study will address these limitations by enlarging the population of thermal anomalies and by improving the methodology to select the fire events to enable the delivery of an accurate segmentation mask.

Finally, it is worth noting that such limitations do not affect the quality of the proposed methodology to automate the design of raw multispectral data but are, instead, specific to the approach used to select target events on L1C for the THRawS dataset, which simply represents a use case of the proposed methodology.

VII. CONCLUSION

This work addressed the lack of available datasets targeting raw optical multispectral data for onboard satellite early detection of target events. To fill this gap, we present the first methodology to automate the creation of dataset containing raw multispectral pushbroom data for target object/event detection. The methodology was demonstrated on Sentinel-2 raw data for the creation of the THRawS dataset, containing warm thermal hotsposts. THRawS features a collection of elements between volcanic, fire events, and Not-events that correspond to more than 1000+ 128×128 patches containing events. To the best of the authors' knowledge, THRawS and the one provided by our previous work [6] are the very first open-source database including raw images from a multispectral pushbroom imager.

The high degree of reproducibility and automation in the application of our methodology is ensured through our open-source toolbox *PyRawS*.

The main motivation behind our work is fostering the research on lightweight preprocessing or end-to-end processing solutions prior to the application of ML processing on board satellites.

TABLE III
DESCRIPTION OF SENTINEL-2 BANDS IN TERMS OF SPATIAL AND SPECTRAL RESOLUTION AND POSSIBLE APPLICATIONS [45]

Band	Spatial resolution [m]	Central Wavelength [nm]	Bandwidth [nm]	Possible applications
B1	60	443 (Ultra Blue)	20	Detection of coastal and aerosol.
B2	10	490 (Blue)	65	Soil and vegetation discrimination, forest mapping, identification of human-made features.
B3	10	560 (Green)	35	Detection of oil or vegetation of water surfaces.
B4	10	665 (Red)	30	Identification of vegetation types, soils, and urban areas.
B5	20	705 (VNIR)	15	Identification of vegetation types.
B6	20	740 (VNIR)	15	Identification of vegetation types.
B7	20	783 (VNIR)	20	Identification of vegetation types.
B8	10	842 (VNIR)	115	Mapping shorelines, biomass content, and detection of vegetation.
B8A	20	865 (VNIR)	20	Identification of vegetation types.
B9	60	940 (SWIR)	20	Detection of water vapor.
B10	60	1375 (SWIR)	30	Cirrus and cloud detection.
B11	20	1610 (SWIR)	90	Measuring moisture content of soil and vegetation, differentiating between snow and clouds.
B12	20	2190 (SWIR)	180	Measuring moisture content of soil and vegetation, differentiating between snow and clouds.

To this aim, for the creation of THRawS, we adopted a lightweight CSC solution for the processing of raw data, which features interesting tradeoffs in terms of computational intensity and quality of registration that makes it promising for its application on board satellites.

Future studies will focus on improving the automation of the proposed methodology and on refining and augmenting the THRawS dataset with a particular emphasis on enhancing the validation approach for label generation. This initiative will involve the rigorous evaluation of label accuracy and the incorporation of advanced techniques to improve the reliability of training data. Critical to this process will be the establishment of robust validation protocols that can effectively handle the inherent complexities and variability of the data sets used. Moreover, we will investigate the suitability of the raw data to be processed by ML on board satellite with minimal pre-processing and compare the obtained trade-offs in terms of application-specific performance and energy efficiency to the current state-of-the-art techniques.

APPENDIX

A. Sentinel-2 Mission

The Sentinel-2 mission is part of the Sentinel constellation and comprises two identical satellites, S2-A and Sentinel-2B (S2-B), which provide data, respectively, since 2015 and 2017. These satellites collectively provide five days of revisit time at the equator, ensuring near-global coverage of land areas, the Mediterranean and closed seas, and coastal waters globally. Each satellite is equipped with a MSI sensor that consists of 12 detectors arranged in a staggered array perpendicular to the flight direction of the satellite. The detectors capture light reflected from the Earth's surface in narrow strips, or "swaths" (290 km width [12]), as the satellite moves forward, which are then combined to create an entire scene.

Each detector measures the Earth's radiance in 13 visible and near-infrared (VNIR) and shortwave infrared (SWIR) bands with a spatial resolution between 10 and 60 m and acquires data

with a pushbroom imaging mode in a sun-synchronous orbit (SSO) [14], [15]. A description of Sentinel-2 bands in terms of spatial resolutions, spectral content and resolutions, and possible target applications is provided in Table III.

B. Coarse Spatial Coregistration Technique

As discussed in Section IV, to design THRawS we leveraged a CSC solution that represents an example of effective lightweight processing that can be applied to raw data. As mentioned in Section VI, the proposed CSC was also applied to a parallel study to perform onboard vessel detection on Raw Sentinel-2 data [9]. This section provides details on the methodology used to design the CSC algorithm and compares it to other B2B alignment solutions in terms of accuracy and latency.

1) CSC Design Methodology: Let us define I the vector of band indices as follows:

$$I \triangleq \{02, 08, 03, 10, 04, 05, 11, 06, 07, 8A, 12, 01, 09\} \quad (8)$$

where I is sorted according to the time delays of the various bands compared to the band B_{02}[46]. Let us consider two bands $B_{I(n)}$ and $B_{I(m)}$, with $B_{I(n)}, B_{I(m)} \in B_S$ where $B_S \triangleq [B_x, B_y, \ldots, B_z]$ is the band collection of a granule that we want to register spatially. To perform the CSC, we apply a shift along and across the satellite-track $\overline{S_{B_{I(n)}\_B_{I(m)}}}$ to the band $B_{I(n)}$ measured with respect to the resolution of $B_{I(n)}$. $\overline{S_{B_{I(n)}\_B_{I(m)}}}$ depends only on the couple of bands $(B_I(n), B_I(m))$, the satellite and the detector number producing that raw data.

For any coupling of satellite, detector number, and couple of bands, the value of $\overline{S_{B_{I(n)}\_B_{I(m)}}}$ can be calculated by using the relations in the following:

$$\begin{cases} \overline{S_{B_{I(n)}\_B_{I(m)}}} = \sum_{k=m}^{n} N_{I(k+1)\_I(k)} \cdot \frac{R_{B_{I(k+1)}}}{R_{B_{I(n)}}}, & n > m \\ \overline{S_{B_{I(n)}\_B_{I(m)}}} = -\overline{S_{B_{I(m)}\_B_{I(n)}}} \cdot \frac{R_{B_{I(n)}}}{R_{B_{I(m)}}}, & n < m \end{cases}$$

$$(9)$$

where $R_{B_{I(k+1)}}$ is the resolution in m of the band $B_{I(k+1)}$, and $N_{I(k+1)\_I(k)}$ is the shift to apply to the band $B_{I(k)}$ to match

the band $B_{I(k+1)}$ measured with respect to the resolution of the band $B_{I(k)}$.

Given a specific coupling, the shift values $N_{I(k+1)\_I(k)}$ are fixed coefficients that we estimated for each of the band couples $(B_{I(k)}, B_{I(k+1)})$ by performing an analysis on volcanic events and Not-events before the filtering of raw data granules. To this aim, we ran a DNN-based method named SuperGlue [43] to extract and match keypoints in a pair of bands $(B_{I(k)}, B_{I(k+1)})$ having the same detector number and used the along-track and across-track distance between the couples of matched key-points to provide an estimation of $N_{I(k+1)\_I(k)}$. More specifically, for a couple of bands $(B_{I(k)}, B_{I(k+1)})$ we proceeded as follows.

First, to ensure that keypoints that were located in adjacent granules could be matched successfully, we first coupled all the granules used for the study that could be stacked along the satellite track. Hence, for a given along-track-stacked granule g, we picked the couple of bands $(B_{I(k)-g}, B_{I(k+1)-g})$. Then, to boost feature presence and distinctiveness without introducing noise in the processing chain, we applied a contrast limited adaptive histogram equalization (CLAHE) [47] algorithm to the stacked bands $(B_{I(k)-g}, B_{I(k+1)-g})$. Then, we applied Super-Glue to extract and match keypoints in the enhanced bands $(B_{I(k)-g}, B_{I(k+1)-g})$ and calculated the average along-track and across-track distance over each couple of matched key-points. We removed the outliers ($\pm 2\sigma$) and averaged the values for both along and across-track offsets. The obtained result provides a value for $N_{I(k+1)\_I(k)-g}$ for the couple of bands $(B_{I(k)-g}, B_{I(k+1)-g})$ of a couple of stacked granules g. Finally, the $N_{I(k+1)\_I(k)}$ is obtained by averaging the $N_{I(k+1)\_I(k)-g}$ over all the couples of stacked granules g. We iterated this process for 13 predefined couples of bands.

Since no granule among the volcano events was generated by detector number 6, to measure $N_{I(k+1)\_I(k)}$, we used the database inclusive of Not-events.

However, we found that band B_{10} was difficult to correlate with the bands B_{03} and B_{04} because of the significant difference in the spectral content. This makes the estimated values of N_{04_10} and N_{10_03} unreliable. To solve this problem, we extracted two additional coefficients N_{10_09} and N_{10_03} and used them to invert (9) and provide better estimates for N_{04_10} and N_{10_03}. The extraction of N_{10_09} was possible because of the higher correlation of the spectral contents of the bands B_{10} and B_{09}.

Regarding settings, we used the superpoint neural network as a feature extractor with outdoor weights as in the original implementation [43]. To increase the detection and matching capabilities, we tuned the hyperparameters for our problem, and the best ones have been reported in our source code[10]. In particular, we increased the matching threshold values to increase the number of strongly matched keypoints, i.e., those having higher probabilities of being the same in both bands.

Concerning the accuracy of the proposed method, Table IV is a good indicator presenting the mean and standard deviation of the across and along-track offsets for the $B_{8\,A} - B_{11}$ spectral

TABLE IV
MEAN AND STANDARD DEVIATION (ALONG- AND ACROSS-TRACK) OF OFFSETS VALUES FOR EACH DETECTOR OF SENTINEL-2 OPTICAL IMAGERS WHEN PERFORMIG THE REGISTRATION OF $B_{8\,A} - B_{11}$ SPECTRAL BANDS

| | Sentinel-2A | | Sentinel-2B | |
D	μ	σ	μ	σ
1	[−174.8, −1.93]	[1.01, 0.26]	[−178.33, −12.78]	[1.12, 0.44]
2	[188.0, −6.0]	[0.76, 0.0]	[186.83, −16.5]	[1.47, 0.55]
3	[−173.08, −2.0]	[0.79, 0.0]	[−174.77, −13.0]	[1.42, 0.0]
4	[186.0, −3.2]	[0.94, 0.42]	[183.22, −14.06]	[0.94, 0.24]
5	[−170.5, −1.88]	[0.53, 0.35]	[−172.5, −13.0]	[0.62, 0.0]
6	[184.0, −2.0]	[0.0, 0.0]	[183.0, −12.75]	[0.82, 0.5]
7	[−170.0, −1.0]	[0.0, 0.0]	[−173.0, −11.67]	[0.0, 0.58]
8	[185.56, −0.64]	[1.12, 0.49]	[184.12, −10.96]	[0.83, 0.2]
9	[−170.55, −1.73]	[1.04, 0.47]	[−172.6, −11.0]	[0.82, 0.0]
10	[187.0, 1.5]	[0.0, 0.58]	[185.13, −9.33]	[0.83, 0.49]
11	[−176.0, −0.5]	[0.0, 0.71]	[−174.88, −11.88]	[1.13, 0.35]
12	[193.0, 4.0]	[0.0, 0.0]	[189.17, −7.33]	[0.41, 0.52]

Here in the offset is provided in pixel.

bands. Such values were obtained by measuring the error due to the proposed CSC method with respect to the shifts that one can obtain using the SuperGlue method. Notably, the standard deviation values exhibit a general trend of being below one pixel for both along- and across-track directions. Although this behavior has been reported for only the $B_{8\,A} - B_{11}$ spectral bands, it is essential to note that this trend has been consistently verified for each spectral band pair of the Sentinel-2 images. Also, note that the results presented in the table have been differentiated based on the satellite and detector number. This differentiation is necessary due to the empirical observation of dependence on the satellite.

Furthermore, an additional interesting pattern that is observable in the results is a change in sign when transitioning from odd to even detectors. This pattern is attributed to the fact that the detectors are inverted on the satellite, resulting in an opposite offset sign for even detectors compared to odd detectors.

It is worth noticing that the reported errors are obtained by using the shift values $N_{I(k+1)\_I(k)}$ matching the average values on the volcanic events before being filtered, which include granules having an acquisition date ranging from 2017 to 2022.

However, in a hypothetical mission scenario, it is possible to recalculate $N_{I(k+1)\_I(k)}$ periodically to optimize them toward specific calibration settings or other possible temporary sensor set-ups.

C. Quality of CSC

In this section, we assess the usability of the proposed CSC for onboard satellite processing by comparing it to a reference algorithm called *SuperGlue* [43] in terms of quality of registration. Such a study was performed across the entire postprocessed THRawS dataset. This comparison is performed by applying SuperGlue and the proposed CSC solution for the alignment of the bands $B_{8\,A}$ and B_{11}. As shown in Table IV, 17 of 24 cases feature a pixel error [σ] smaller than a pixel for both the along and across-track directions. In three cases, the coregistration error is null. The maximum along-track and across-track errors are 1.47 px and 0.71 px, respectively. Such values are significantly lower than the smallest thermal anomaly in the THRawS dataset,

TABLE V
Evaluation of Timing Performance for SuperGlue, LightGlue (ALIKED, DISK, SIFT, Superpoint), and Traditional SIFT-Based Method in Spatial Coregistration Compared to the Developed CSC Approach, Executed on CPU (Intel Ⓡ Xeon Ⓡ Gold 6248 CPU @ 2.50 GHz) and GPU (NVIDIA A40-48 C With Driver Version 525.105.17 and CUDA Version 12.0)

Method	Average CPU [ms]	Average GPU [ms]
SuperGlue (Superpoint)	5167.45	1706.30
Lightglue (ALIKED)	5651.29	N/A
Lightglue (DISK)	7169.25	536.51
Lightglue (SIFT)	1916.15	1225.58
Lightglue (Superpoint)	6660.91	458.92
SIFT	23849.50	2189.06
CSC	**16.63**	**1.65**

which was selected using the nine contiguous pixel criterion implemented by Massimetti et al. to suppress false positives [29]. Therefore, no warm temperature hotspot is misaligned due to the coregistration error.

These results showcase the potential of the proposed CSC algorithm, not for its precision in B2B alignment but for its energy efficiency and simplification of the problem. The primary aim of CSC is to correct misalignments rather than performing complex warping. We juxtapose the CSC method with other techniques, noting that while others rely on feature detection and matching, CSC efficiently bypasses these steps.

1) CSC Timing Performance: This section compares the proposed CSC technique in terms of timing performance to Super-Glue, and other coregistration solutions including the traditional SIFT implementation [34], LightGlue [44] method, covering different feature descriptors, including Superpoint, ALIKED [48], and DISK [49]. Differently from the described CSC, these techniques are feature-based, i.e., they extract and match keypoints from the bands to be aligned enabling a fine B2B alignment. Results shown in Table V showcase the profiling of these methods. This comparison was conducted across the granules of the THRawS dataset, highlighting the practical advantages of CSC in onboard satellite processing contexts.

More specifically, we set up an experiment to compare the time to register ten raw data granules.

For all the aforementioned methods, we measured the time to register the bands B_{8A}, B_{11}, and B_{12}. We profiled both the methods by using PyTorch Profiler [50] both on an Intel Ⓡ Xeon Ⓡ Gold 6248 CPU @ 2.50 GHz central processing unit (CPU) and an NVIDIA A40-48 C graphics processing unit (GPU) to test the dependency on the time performance on the hardware device. Each test case was run three times, and time performance was measured by averaging the results on the three tests. All the tests were performed continuously with 15 warm-up cycles at the beginning of the test series with one raw data granule.

Concerning the CSC, precalculated shifts were stored in memory before running the tests to reduce the overhead time due to the storage memory access. When the other methods were used, shift values were retrieved using the approach and the model set-up described in Section B.

On both devices, all the methods show a linear increment of the time to perform the band registration with the number of images. In general, on both CPU and GPU, our method outperforms the other feature-matching techniques. In particular, for a single granule with bands B_{8A}, B_{11}, and B_{12} our methodology requires 17 ms and 1.65 ms on CPU and GPU, respectively.

DATA AVAILABILITY

The dataset associated with this study is available and can be accessed through the following DOI: 10.5281/zenodo.8367083.

REFERENCES

[1] G. Furano et al., "Towards the use of artificial intelligence on the edge in space systems: Challenges and opportunities," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 35, no. 12, pp. 44–56, Dec. 2020.

[2] G. Giuffrida et al., "The ϕ-sat-1 mission: The first on-board deep neural network demonstrator for satellite earth observation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5517414.

[3] G. Guerrisi, F. D. Frate, and G. Schiavon, "Convolutional autoencoder algorithm for on-board image compression," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 151–154.

[4] G. Guerrisi, F. D. Frate, and G. Schiavon, "Artificial intelligence based on-board image compression for the ϕ-sat-2 mission," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 8063–8075, 2023.

[5] P. D. Stasio, A. Sebastianelli, G. Meoni, and S. L. Ullo, "Early detection of volcanic eruption through artificial intelligence on board," in *Proc. IEEE Int. Conf. Metrol. Extended Reality Artif. Intell. Neural Eng. (MetroXRAINE)*, 2022, pp. 714–718.

[6] M. P. D. Rosso, A. Sebastianelli, D. Spiller, P. P. Mathieu, and S. L. Ullo, "On-board volcanic eruption detection through CNNs and satellite multispectral imagery," *Remote Sens.*, vol. 13, no. 17, 2021, Art. no. 3479.

[7] V. Růžička et al., "RaVÆn: Unsupervised change detection of extreme events using ML on-board satellites," *Sci. Rep.*, vol. 12, no. 1, Oct. 2022, Art. no. 16939. [Online]. Available: https://www.nature.com/articles/s41598-022-19437-5

[8] G. M.-Garcia et al., "Towards global flood mapping onboard low cost satellites with machine learning," *Sci. Rep.*, vol. 11, no. 1, Mar. 2021, Art. no. 7249. [Online]. Available: https://www.nature.com/articles/s41598-021-86650-z

[9] R. D. Prete, G. Meoni, N. Longépé, M. D. Graziano, and A. Renga, "First results of vessel detection with onboard processing of Sentinel-2 raw data by deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2023, pp. 6262–6265.

[10] V. Fanizza, D. Rijlaarsdam, P. T. T. González, and J. L. E.-Aranda, "Transfer learning for on-orbit ship segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 21–36.

[11] K. Thangavel, D. Spiller, R. Sabatini, P. Marzocca, and M. Esposito, "Near real-time wildfire management using distributed satellite system," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 8063–8075, 2023.

[12] M. Drusch et al., "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425712000636

[13] M. A. Wulder et al., "Current status of landsat program, science, and applications," *Remote Sens. Environ.*, vol. 225, pp. 127–147, 2019.

[14] J. A. Richards and J. A. Richards, *Remote Sensing Digital Image Analysis*, vol. 5. Cham, Switzerland: Springer, 2022.

[15] "*Sentinel-2 Products Specification Document*," Accessed: Jan. 21, 2023. [Online]. Available: https://sentinel.esa.int/documents/247904/685211/sentinel-2-products-specification-document

[16] "Sentinel online - calibration," Accessed: Mar. 4, 2023. [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/calibration

[17] A. S. Danielsen, T. A. Johansen, and J. L. Garrett, "Self-organizing maps for clustering hyperspectral images on-board a cubesat," *Remote Sens.*, vol. 13, no. 20, 2021, Art. no. 4174.

[18] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 7, pp. 2217–2226, Jul. 2019.

[19] M. B. Henriksen, J. L. Garrett, E. F. Prentice, A. Stahl, T. A. Johansen, and F. Sigernes, "Real-time corrections for a low-cost hyperspectral instrument," in *Proc. 10th Workshop Hyperspectral Imag. Signal Process. Evol. Remote Sens. (WHISPERS)*, 2019, pp. 1–5.

[20] N. Melega et al., "Implementation of the ϕsat-2 on board image processing chain," in *Proc. Sensors, Systems, and Next-Generation Satellites XXVII*, SPIE, 2023, pp. 264–276.

[21] G. M.-García et al., "In-orbit demonstration of a re-trainable machine learning payload for processing optical imagery," *Sci. Rep.*, vol. 13, no. 1, 2023, Art. no. 10391.

[22] D. Derksen, G. Meoni, G. Lecuyer, A. Mergy, M. Märtens, and D. Izzo, "Few-shot image classification challenge on-board," in *Proc. Workshop-Data Centric AI NeurIPS*, 2021.

[23] G. Meoni et al., "The OPS-SAT case: A data-centric competition for onboard satellite image classification," *Astrodynamics*, pp. 1–22, 2024, doi: 10.1007/s42064-023-0196-y.

[24] "AI on the edge: "the OPS-SAT case"," Accessed: Mar. 6, 2023. [Online]. Available: https://kelvins.esa.int/opssat

[25] "OPS-SAT," Accessed: Feb. 20, 2023. [Online]. Available: https://www.esa.int/Enabling_Support/Operations/OPS-SAT

[26] R. Kennedy and W. Cohen, "Automated designation of tie-points for image-to-image coregistration," *Int. J. Remote Sens.*, vol. 24, no. 17, pp. 3467–3490, 2003.

[27] E. Venzke, "Global volcanism program," [Database] Volcanoes of the World (v. 5.0.1; 19 Dec.2022), Distributed by Smithsonian Institution, 2022. Accessed: Jul. 5, 2024. [Online]. Available: https://doi.org/10.5479/si.GVP.VOTW5-2024.5.2

[28] P. L. Abbott, *Natural Disasters*, 6th ed. New York, NY, USA: McGraw-Hill Higher Educ., 2008.

[29] 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 Sens.*, vol. 12, no. 5, 2020, Art. no. 820. [Online]. Available: https://www.mdpi.com/2072-4292/12/5/820

[30] P. E. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transform techniques," *IEEE Trans. Geosci. Electron.*, vol. 8, pp. 353–368, Oct. 1970.

[31] P. Nandy, B. N. Post, J. L. Smith, and M. L. Decker, "Edge-based correlation image registration algorithm for the multispectral thermal imager (MTI)," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery*, vol. 5425. Bellingham, WA, USA: SPIE, 2004, pp. 97–207.

[32] G. Yu, T. Vladimirova, and M. Sweeting, "Autonomous band registration for on-board applications," in *Proc. IEEE Int. Conf. Signal Process. Commun.*, 2007, pp. 1327–1330.

[33] A. Ordóñez, D. B. Heras, and F. Argüello, "Comparing area-based and feature-based methods for co-registration of multispectral bands on GPU," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 1575–1578.

[34] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[35] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," *Comput. Vis. Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[36] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 2548–2555.

[37] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 2564–2571.

[38] "scikit-image," Accessed: Jan. 31, 2023. [Online]. Available: https://scikit-image.org/

[39] Y. Liu, W. Zhi, B. Xu, W. Xu, and W. Wu, "Detecting high-temperature anomalies from Sentinel-2 MSI images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 177, pp. 174–193, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271621001337

[40] N. Andela, J. W. Kaiser, G. R. v. d. Werf, and M. J. Wooster, "New fire diurnal cycle characterizations to improve fire radiative energy assessments made from modis observations," *Atmospheric Chem. Phys.*, vol. 15, no. 15, pp. 8831–8846, 2015. [Online]. Available: https://acp.copernicus.org/articles/15/8831/2015/

[41] A. Sebastianelli, M. P. D. Rosso, and S. L. Ullo, "Automatic dataset builder for machine learning applications to satellite imagery," *SoftwareX*, vol. 15, 2021, Art. no. 100739.

[42] "Sentinel-2 calibration and validation plan for the operational phase," Accessed: Aug. 31, 2023. [Online]. Available https://sentinel.esa.int/documents/247904/2047089/sentinel-2_cal-val_phase-e2

[43] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superglue: Learning feature matching with graph neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 4938–4947.

[44] P. Lindenberger, P. E. Sarlin, and M. Pollefeys, "Lightglue: Local feature matching at light speed," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2023, pp. 17627–17638.

[45] "Sentinel2hub - sentinel-2 bands," Accessed: Nov. 9, 2023. [Online]. Available: https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/bands/

[46] R. Binet, E. Bergsma, and V. Poulain, "Accurate Sentinel-2 inter-band time delays," *ISPRS Ann. Photogrammetry Remote Sens. Spatial Inf. Sci.*, vol. 1, pp. 57–66, 2022.

[47] P. Musa, F. A. Rafi, and M. Lamsani, "A review: Contrast-limited adaptive histogram equalization (CLAHE) methods to help the application of face recognition," in *Proc. 3rd Int. Conf. Informat. Comput.*, 2018, pp. 1–6.

[48] X. Zhao, X. Wu, W. Chen, P. C. Chen, Q. Xu, and Z. Li, "Aliked: A lighter keypoint and descriptor extraction network via deformable transformation," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 5014016.

[49] M. Tyszkiewicz, P. Fua, and E. Trulls, "Disk: Learning local features with policy gradient," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 14254–14265, 2020.

[50] "PyTorch Profiler," Accessed: Feb. 13, 2023. [Online]. Available: https://pytorch.org/tutorials/recipes/recipes/profiler_recipe.html

Gabriele Meoni received the M.Sc. degree in electronic engineering and the Ph.D. degree in information engineering, under the supervision of Prof. Luca Fanucci, from the University of Pisa, Pisa, Italy, in 2016 and 2020, respectively.

After completing his doctoral studies, from 2020 to 2023 he was a Internal Research Fellow with the European Space Agency (ESA) (Advanced Concepts Team (ACT) 2020–2021 and Φ-lab 2021–2023), where he conducted research on artificial intelligence (AI) and neuromorphic computing for onboard spacecraft applications. During 2022–2023, he was a visiting Researcher with AI Sweden, focusing on distributed edge learning for satellite constellations. From 2023 to 2024, he was an Assistant Professor with the Faculty of Aerospace Engineering at Delft University of Technology. He is currently an Innovation Officer with ESA, Frascati, Italy. He has coauthored more than 40 scientific publications. His research interests include satellite onboard processing, AI for Earth observation, and neuromorphic computing.

Roberto Del Prete was born in 1994. He received the bachelor's and master's degrees in aerospace engineering in 2017 and 2021, respectively from the University of Naples Federico II, Naples, Italy, where he is currently working toward the Ph.D. degree in onboard AI for remote sensing systems and autonomous navigation.

He is an Italian Researcher specializing in deep learning and edge computing for remote sensing. His notable work includes the development of "FederNet," a terrain relative navigation system. His professional experience includes roles as a Visiting Researcher with the European Space Agency's Φ-Lab and SmartSat CRC in Australia. He has authored or coauthored 28 scientific publications and 16 conference presentations, dedicated to leveraging advanced technologies to address global challenges in remote sensing and AI. His research focuses on improving time-critical decision-making through advanced AI solutions for space missions and Earth monitoring.

Mr. Prete was the recipient of the 2022 NATO STO Early Career Scientist Award and participated in the 2021 NASA-ESA Trans-Atlantic Training. He has contributed to key projects, such as Kanyini Mission, and PYRAWS, developing AI algorithms for real-time maritime monitoring and thermal anomaly detection.

Federico Serva received the M.Sc. degree in physics from the University of Rome Tor Vergata, Rome, Italy, in 2014, and the Ph.D. degree in environmental sciences from the University of Naples Parthenope, Naples, Italy, in 2018.

After the Ph.D., he has been a postdoc with the National Research Council, working with climate model and Earth observation data for model evaluation and process studies, and a joint postdoc between the Italian and European Space Agencies. He is currently a fixed-term Researcher with the National Research Council and a visiting Researcher with the ESA Φ-Lab. His research interests include the hyperspectral imaging data processing and applications of machine learning to weather and climate forecasting.

Alix De Beusscher received the M.Sc. degree in data science from UCLouvain, Belgium, in 2021.

Alix completed a Young Graduate Trainee (YGT) program with the European Space Agency (ESA). Alix is currently an Aerospace Engineer with Aerospacelab, Mont-Saint-Guibert, Belgium. Alix's research interests include AI for data insights to satellite system engineering, leveraging a unique combination of data science and aerospace engineering skills to advance technology and research in the aerospace sector.

Olivier Colin received the M.Sc. degree in computer networks from University of Surrey, Guildford, U.K.

Since 2009, Olivier Colin has been working with the European Space Agency (ESA). He is currently with the Copernicus and Ground Segment / Data Management Division, ESA, Frascati, Italy.

Nicolas Longépé received the M.Eng. degree in electronics and communication systems and the M.Sc. degree in electronics from the National Institute for the Applied Sciences, Rennes, France, in 2005, and the Ph.D. degree in signal processing and telecommunication from the University of Rennes I, Rennes, in 2008.

From 2007 to 2010, he was with the Earth Observation Research Center, Japan Aerospace Exploration Agency, Tsukuba, Japan. From 2010 to 2020, he was a Research Engineer with the Space Observation Division, Collecte Localization Satellites, Plouzané, France. Since September 2020, he has been an Earth observation data Scientist, Φ−Lab Explore Office, European Space Research Institute, European Space Agency, Frascati, Italy. He has been working on the development of innovative synthetic aperture radar-based applications for environmental and natural resource management (ocean, mangrove, land and forest cover, soil moisture, snow cover, and permafrost) and maritime security (oil spills, sea ice, icebergs, and ship detection/tracking). At the Φ−Lab, he is particularly involved in the development of innovative Earth observation missions in which artificial intelligence is directly deployed at the edge (on the spacecraft). His research interests include Earth observation remote sensing and digital technologies, such as machine (deep) learning.