Sen4Map: Advancing Mapping With Sentinel-2 by Providing Detailed Semantic Descriptions and Customizable Land-Use and Land-Cover Data

Surbhi Sharma, *Student Member, IEEE*, Rocco Sedona, *Member, IEEE*, Morris Riedel, *Member, IEEE*, Gabriele Cavallaro, *Senior Member, IEEE*, and Claudia Paris, *Senior Member, IEEE*

***Abstract*—This article presents Sen4Map, a large-scale benchmark dataset designed to enhance the capability of generating land-cover maps using Sentinel-2 data. Comprising nonoverlapping 64 × 64 patches extracted from Sentinel-2 time series images, the dataset spans 335 125 geotagged locations across the European Union. These locations are associated with detailed land-cover and land-use information gathered by expert surveyors in 2018. Unlike most existing large datasets available in the literature, the presented database provides: first, a detailed description of the land-cover and land-use properties of each sampled area; second, independence of scale, as it is associated with reference data collected in situ by expert surveyors; third, the ability to test both temporal and spatial classification approaches because of the availability of time series of 64 × 64 patches associated with each labeled sample; and fourth, samples were collected following a stratified random sample design to obtain a statistically representative spatial distribution of land-cover classes throughout the European Union. To showcase the properties and challenges offered by Sen4Map, we benchmarked the current state-of-the-art land-cover classification approaches.**

***Index Terms*—Benchmark dataset, land-use and land-cover mapping, machine learning, Sentinel-2, supervised classification, Land Use and Coverage Area frame Survey (LUCAS).**

I. Introduction

THE expanded deployment of Earth observation (EO) satellites produces vast quantities of remote sensing (RS) data in multiple modalities, such as optical and radar, with resolutions from submetric to decametric. Extracting meaningful information from these varied datasets is crucial for effectively addressing environmental challenges such as urbanization, deforestation, climate change, and global warming [1]. This extraction provides important insights into environmental dynamics and supports sustainable decision-making processes.

Classification maps, produced through the analysis of Remote Sensing (RS) data, stand as key thematic outputs. They offer crucial insights into the Earth's physical and functional characteristics by distinguishing between natural landscapes and human activities across various time periods. They differentiate land-cover, which describes physical types such as corn fields or grasslands, from land-use, revealing how these areas are employed by humans, for instance, for agriculture. Given the vast amounts of data from satellite image time series, developing automatic methods for producing classification maps has become essential [2].

The introduction of the Copernicus Programme, providing free and open access to RS data, along with advancements in machine learning (ML), has revolutionized Earth's surface monitoring. This combination has significantly improved the accuracy of land-cover map generation [3]. The Sentinel-2 constellation, part of this program, acquires multispectral images having 13 spectral bands, out of which four bands at a spatial resolution of 10 m, six bands at 20 m, and three bands at 60 m. The Sentinel-2 constellation has a temporal resolution of five days [4]. These satellite image time series, covering the global surface with short revisit periods, provide the foundation for large-scale mapping such as detailed crop type mapping [5], forest monitoring [6], water body detection [7], and the production of global thematic products at a 10 m spatial resolution [8], [9].

To facilitate the development of workflows based on novel deep learning (DL) algorithms and the evaluation and comparison of classical Machine Learning (ML) approaches, several large-scale datasets based on Sentinel-2 data have been released. These benchmarks address different tasks, such as multisensor and multimodal data fusion [10], cloud removal [11], [12], super-resolution [13], visual question answering [14], as well as large-scale datasets for unsupervised pretraining and transfer learning [15], [16], [17], [18]. The availability of web platforms such as HyperLableMe [19] also facilitates benchmarking of RS image classifiers with their harmonized large-scale dataset containing labeled and heterogeneous multispectral and hyperspectral images.

Manuscript received 26 March 2024; revised 3 July 2024; accepted 22 July 2024. Date of publication 29 July 2024; date of current version 15 August 2024. This work was supported in part by the ADMIRE Project, which received funding from the European Union's Horizon 2020 JTI-EuroHPC research and innovation programme under Grant 956748, in part by the German Federal Ministry of Education and Research under Grant 16HPC008, and in part by the EUROCC2 project funded by the European High-Performance Computing Joint Undertaking (JU) and EU/EEA states under Grant 101101903. *(Corresponding author: Gabriele Cavallaro.)*

Surbhi Sharma, Morris Riedel, and Gabriele Cavallaro are with the Jülich Supercomputing Centre, Wilhelm-Johnen-Straße, 52428 Jülich, Germany, and also with the University of Iceland, 107 Reykjavik, Iceland (e-mail: su.sharma@fz-juelich.de; morris@hi.is; g.cavallaro@fz-juelich.de).

Rocco Sedona is with the Jülich Supercomputing Centre, Wilhelm-Johnen-Straße, 52428 Jülich, Germany (e-mail: r.sedona@fz-juelich.de).

Claudia Paris is with the University of Twente, 7522 NB Enschede, The Netherlands (e-mail: c.paris@utwente.nl).

The dataset and code can be downloaded at: https://datapub.fz-juelich.de/sen4map.

Digital Object Identifier 10.1109/JSTARS.2024.3435081

The release of Earth Observation (EO) benchmarks is an essential trend in establishing a large scale, open, and potentially interoperable ecosystem of open-access datasets. This trend is vital for keeping pace with the escalating demands of large-scale Deep Learning (DL) models, such as foundation models, which are increasingly data-intensive [20]. DL models for EO necessitate substantial resources for collecting and compiling the datasets required for training. Providing the community with large-scale land-cover and land-use datasets not only offers an immediately beneficial resource but also acts as a model for future expansions, facilitating significant progress in mapping the Earth's surface [18]. Within the domain of land-cover classification, datasets such as BigEarthNet [21], EuroSAT [22], Worldcover [23], and Dynamic World [8] are prominent. Worldcover is an European Space Agency (ESA) product at the global level, recently generated using Sentinel-1 and Sentinel-2 images by training a random forest (RF) classifier with a labeled dataset of 141 000 training samples generated through manual annotation of Google Maps and Microsoft Bing images. The map legend includes 11 widely spread generic classes and also has been used to generate the global Google's Dynamic World [8] and Esri's [24] land-cover products. Unlike the ESA Worldcover Map, these land-cover products were generated using convolutional DL models trained with labeled data derived from manual interpretation of Sentinel-2 images. A high spatial resolution of the CORINE land-cover (CLC) map, the CLC+, has been recently released [25] to provide the first European land-cover product in unprecedented detail, leveraging Sentinel-1 and Sentinel-2 images acquired in 2018. Based on the temporal convolutional DL model, the map aims to depict 11 land-cover categories as presented in [26]. On the other hand, using these benchmark datasets can present challenges from an operation viewpoint. They usually feature a legend with fewer than 15 land-cover classes, reducing the level of detail in Earth's surface descriptions and oversimplifying the complexity and diversity of land-cover types. When annotations are performed through photo-interpretation, labeled datasets heavily rely on the EO data used. In addition, their specific map legends may render them inadequate for testing different classification schemes or improving the semantic detail a land-cover map can offer [27].

This article presents a benchmark dataset, Sen4Map, which is designed to advance the state-of-the-art of land-cover mapping applied to Sentinel-2 satellite data. Differently from the existing datasets, Sen4Map is the first large-scale dataset to include 335 125 time series of 64×64 multispectral patches extracted from Sentinel-2 tiles, each associated with geolocated land-cover data collected in situ by experienced surveyors [28], [29]. To this end, we leveraged the public availability of the Land Use and Coverage Area frame Survey (LUCAS) database, a triennial ground survey coordinated by the Statistical Office of the European Commission (Eurostat) that collects land-cover and land-use data over the whole European Union [29] since 2001. The LUCAS data have already been used to validate existing thematic products [30], [31], as well as to generate land-cover maps at country [32] or European scale [33]. However, all these approaches have made partial use of the extremely valuable information provided by the LUCAS database, generally considering only high-level land-cover categories (e.g., cropland, artificial surfaces, and water bodies, etc.) and completely neglecting land-use categories. In addition, most approaches using Sentinel-2 data [32], [33] or Landsat data [34] considered only the LUCAS points associated with homogeneous land-cover, by filtering out a high amount of samples that may be difficult to classify.

In this article, we aim to enhance the capability of Sentinel-2 for land-cover mapping by fully leveraging the challenges presented by the real in situ LUCAS data collection, which provides land-cover information not specifically tailored to the properties of Sentinel-2 satellite data. Specifically, the main contributions of the proposed reference dataset are as follows.

1) The dataset provides detailed, geotagged land-cover and land-use information associated with the central pixel of each 64×64 patch extracted from Sentinel-2 tiles, enabling an assessment of the limitations of satellite data in producing semantically detailed land-cover maps.
2) The scale-independent land-cover information, collected in-situ by experienced surveyors, facilitates the analysis of Sentinel-2's ability to accurately map land-covers of various sizes.
3) The provision of ready-to-use 64×64 Sentinel-2 time series data supports evaluating and comparing land-cover mapping algorithms that utilize both spatial and temporal dimensions.
4) Sen4Map's design, based on a stratified random sampling across the European Union, assesses the generalization capabilities of land-cover mapping techniques, both within individual countries and across the EU.

The rest of this article is organized into five sections. Section II provides an overview of the proposed approach for developing the Sen4Map dataset. Section III describes the information included in the datasets, including spatial and temporal properties. Section IV discusses the experimental results obtained by state-of-the-art classification algorithms and highlights the properties and challenges presented by Sen4Map. Finally, Section V concludes this article and presents potential future developments.

II. Sen4Map Generation

This section outlines the methodology for developing the Sen4Map, leveraging two data sources: the 2018 Sentinel-2 image time series and the 2018 LUCAS field data. Subsequent sections provide further details.

A. LUCAS Database

Sen4Map includes geotagged locations of the LUCAS points, which provide information on land-cover (the physical coverage observed on the Earth's surface) and land-use (the socio-economic function of the observed Earth's surface). Moreover, for each LUCAS point, we can link to information that is available in the dataset such as several environmental parameters, a downward-facing photo of the target point (P), and four landscape photos facing the cardinal directions (N, E,

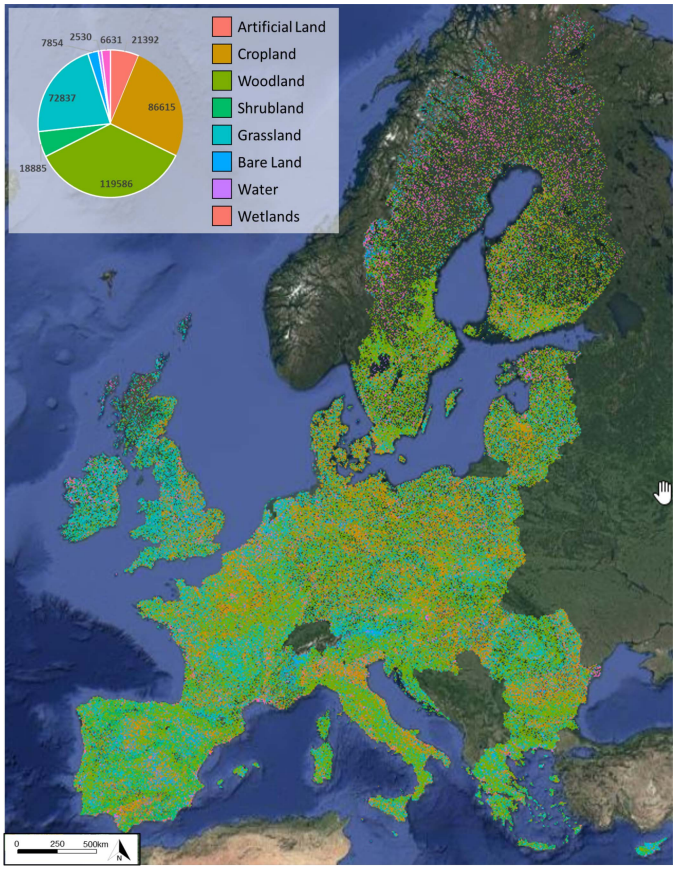


Fig. 1. Spatial distribution of the 2018 LUCAS samples collected all over the European Union. The different land-cover categories are highlighted with different colors.

S, W), as described in [29]. These data, along with detailed metadata concerning the quality of the observations, including the acquisition date, the GPS location of the surveyor in the field, and the theoretical location of the target point (used as the central pixel of the 64×64 patch), are publicly available. Given the recent launch of Sentinel-2 in August 2015, the initial version of Sen4Map primarily utilizes the consolidated 2018 LUCAS database. This approach facilitates the creation of an annual time series of contemporary Sentinel-2 images for each LUCAS point. Nevertheless, our goal is to regularly update the Sen4Map dataset with recent Sentinel-2 satellite acquisitions following the release of the consolidated 2022 LUCAS dataset. The spatial distribution of the 2018 LUCAS samples collected all over the European Union is shown in Fig. 1 by representing the main land-cover classes in different colors.

B. Sentinel-2 Database

Sen4Map is based on multispectral time series of patches, each 64×64 pixels, extracted from Sentinel-2 tiles of 100×100 km, acquired in 2018 by the Sentinel-2 A and Sentinel-2B satellites. Launched in June 2015 and March 2017, respectively, these land-monitoring satellites operate in a sun-synchronous orbit. They are both equipped with a Multispectral Instrument which can capture 13 spectral bands. The bands B01 (Aerosols),

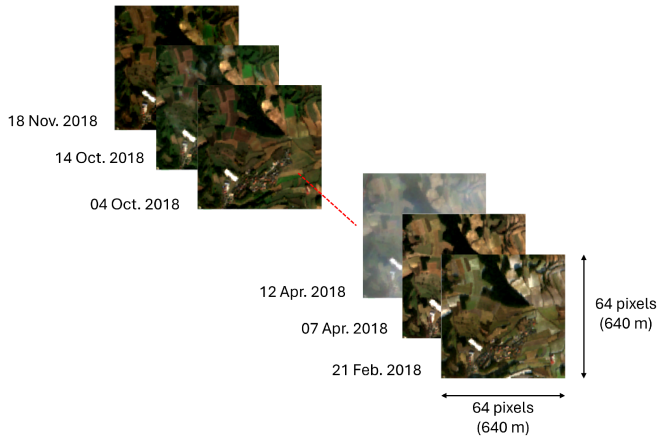


Fig. 2. Example of a Sentinel-2 time series from the Sen4Map benchmark dataset associated with one LUCAS point. In this qualitative example, only the images acquired from the 21st of February until the 18th of November are visualized to capture the temporal component that can characterize different land cover classes.

B09 (Water Vapour), and B10 (Cirrus) feature a spatial resolution of 60 m and central wavelengths of 443 nm, 945 nm, and 1375 nm, respectively. Bands B05 (Red Edge 1), B06 (Red Edge 2), B07 (Red Edge 3), B08 A (Red Edge 4), B11 (SWIR 1), and B12 (SWIR 2) offer a spatial resolution of 20 m with central wavelengths of 705 nm, 740 nm, 783 nm, 865 nm, 1610 nm, and 2190 nm, respectively. Finally, bands B02 (Blue), B03 (Green), B04 (Red), and B08 (NIR) provide a spatial resolution of 10 m with central wavelengths of 490 nm, 560 nm, 665 nm, and 842 nm, respectively. The Sen4Map dataset provides the Sentinel-2 bands only from 10 m and 20 m spatial resolution since these bands are typically used for land-cover mapping. The bands with 60 m spatial resolution such as Aerosols, Water Vapour, and Cirrus are not included in the dataset, as they are primarily intended for atmospheric corrections. Due to the short repeat cycle for the two-satellite constellation (Sentinel-2 A and Sentinel-2B) of approximately 5 days, the time series images obtained are ideal for continuous mapping and monitoring of the Earth's surface. Fig. 2 shows a qualitative example of Sentinel-2 time series reported for a 64×64 patch, where only the true color compositions are represented per acquisition.

III. Sen4Map Properties

A. Spatio-Temporal Properties

The Sentinel-2 tiles considered for the extraction of a time series of patches were those with cloud coverage of less than 40% and with less than 5% no-pixel data. We selected the locations where LUCAS points are available (i.e., theoretical locations, 2006–2018), with the support of Google Earth Engine (GEE) [35]. LUCAS is carried out systematically using a distributed field design based on a $2\,\text{km} \times 2\,\text{km}$ grid throughout the European Union. The points of intersection from the grid are referred to as theoretical locations, "*th_long*" and "*th_lat*" [29]. This stratified spatial distribution allows for properly mapping the different land-cover conditions while reducing the

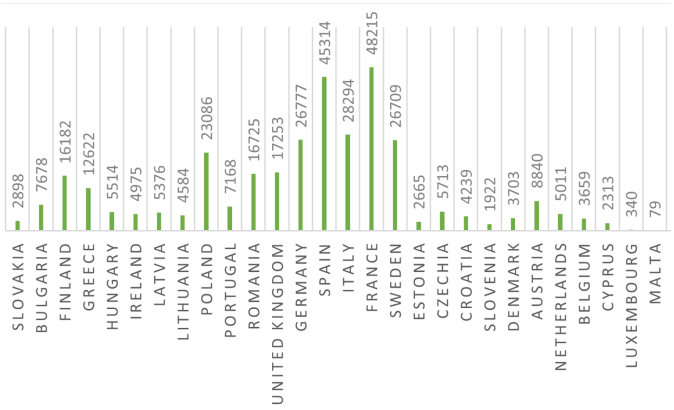


Fig. 3. Country-wise distribution of LUCAS surveyed locations with available in-situ data.

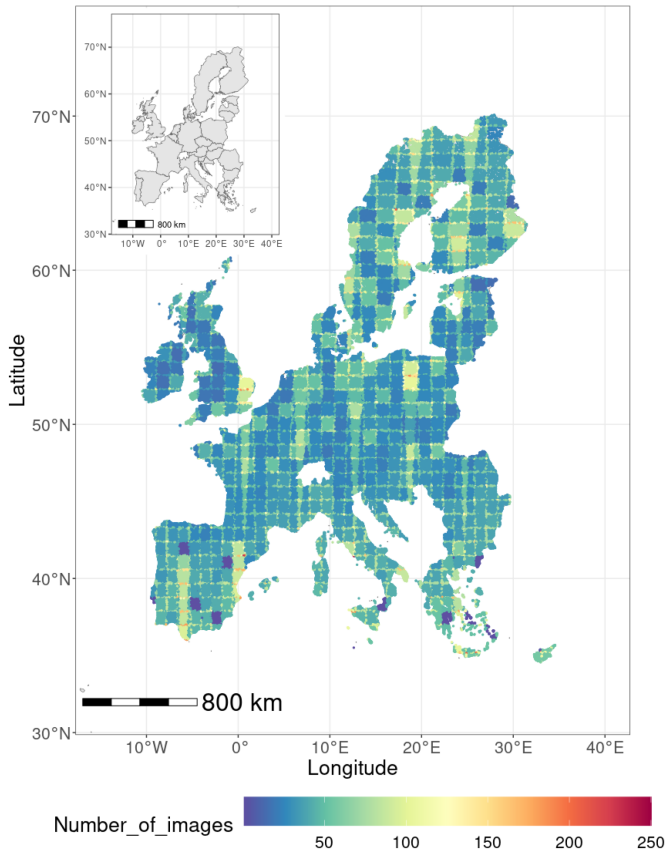


Fig. 4. Number of Sentinel-2 images having less than 40% of cloud coverage and associated with each LUCAS point.

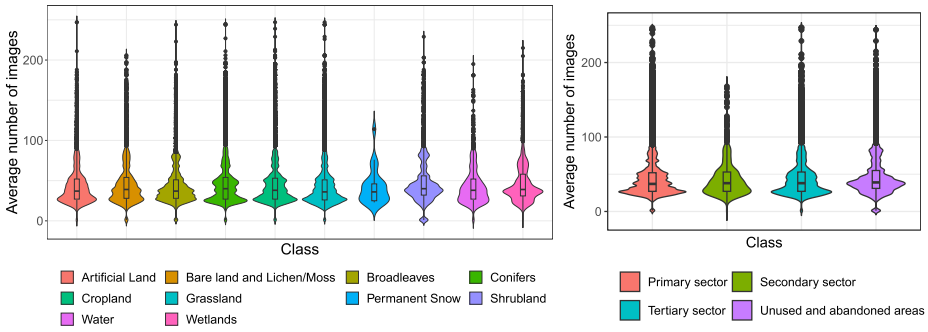
risk of spatial autocorrelation, which is extremely important for land-cover mapping. Although, the field survey is carried out based on theoretical locations, the discrepancy between the actual surveyed location,"gps_lon" and "gps_lat" and the theoretical locations, could also induce spatial errors. A buffer around the point can therefore help mitigate this based on in situ observations of the multiyear harmonized LUCAS survey database [29]. We created a $650\,\mathrm{m}$ buffer around each point to extract patches of 64×64 size. The choice of a 64×64 patch is critical for various ML algorithms. Since several Sentinel-2 tiles in the GEE image collection are rotated by differing orbits, there is a need to compensate for NO DATA at the edges, requiring the buffer size of $650\,\mathrm{m}$. This results in some patches being larger than 64×64, which are subsequently cropped to the target size. Each time series includes additional information on scene classification maps "SCL" at a spatial resolution of $20\,\mathrm{m}$, along with spatial resolutions of 10 and $20\,\mathrm{m}$.

Out of a total of 337 854 LUCAS points from the year 2018, only 2 704 points had no Sentinel-2 times series images considering less than 40% of cloud coverage and less than 5% of no-pixel data. Another 25 LUCAS points where no in situ data were available (i.e., no labels) were also discarded. Therefore, the Sen4Map dataset results in a total of 335 125 time-series patches across 28 European countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Ireland, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom. The distribution of the LUCAS points all over the European Union is shown in Fig. 3 by presenting per country the number of geotagged locations where in-situ data are available. The average number of Sentinel-2 images available per time-series for the respective LUCAS locations are reported in Fig. 4. We considered only Sentinel-2 tiles with cloud coverage lower than 40%, leading to time series of varying lengths and temporal sampling rates across different LUCAS locations. Spatial and temporal harmonization of satellite data is a well-known challenge in large-scale land-cover mapping. Various algorithms for temporal feature sampling [36], [37] can be employed to mitigate differences in the acquisition times of the first and last images, as well as the diverse temporal sampling and sequence lengths associated with each LUCAS point time series.

B. Land-Cover and Land-Use Information

For each LUCAS point, a hierarchical representation of the associated land-cover and land-use is reported. The land-cover types reported in the first level provide a high-level overview of what is present in the scene, by distinguishing "Artificial land" (A00), "Cropland" (B00), "Woodland" (C00), "Shrubland" (D00), "Grassland" (E00), "Bareland and lichens/moss" (F00), "Water areas" (G00), and "Wetlands" (H00). Similarly, the first level of land-use type aims at classifying "Primary sector" (U100), "Secondary sector" (U200), "Tertiary sector transport, utilities and residential" (U300), and "Unused and abandoned areas" (U400). The statistical distribution of the number of Sentinel-2 images associated with higher level land-cover and land-use classes is shown in Fig. 5. The detailed countrywise statistics of the Sen4Map dataset for the higher level hierarchy of LUCAS labels are shown in the Appendix.

The land-cover labels are defined according to the land-cover classification system (LCCS) [38], a comprehensive reference



(a) (b)

Fig. 5. Statistical distribution depicted as a violin plot of the number of Sentinel-2 images associated with each (a) land-cover class and (b) land-use class. The first level of the hierarchy is reported.

system designed for land-cover mapping. The land-cover Classification System (LCCS) aims to provide extensive flexibility, enabling the description of land-cover features at various scales or levels of detail. Rather than offering a single, universal thematic legend, it standardizes terminology and attributes used to define land-cover classes. The system utilizes independent diagnostic criteria, facilitating correlation with existing classifications and legends regardless of the mapping scale or methods used.

The descriptions of the land-cover **Lc1** and land-use **Lu1** types are very detailed in the second level of the hierarchy, where the specific combination of numbers and digits provides a detailed definition of what is present in the scene, e.g., the "A11" category refers to "Buildings with one to three floors." Similarly, "U112" is associated with fallow land to describe when the land is grazed and not used for crop production. In this framework, intersecting the information provided by the land-use and land-cover descriptions allows for an extremely detailed characterization of the geotagged location. For instance, parking areas close to an area used for forestry purposes will have a land-cover label "A21" (non-built-up area features) and land-use "U120" (areas used for forestry purposes). Moreover, for patches where more than one land-cover coexists, a second land-cover **Lc2** and land-use **Lu2** information are reported together with their percentages of coverage.

C. Structure of the Dataset

The Sen4Map dataset is available for 28 countries in individual HDF5 files. The HDF5 format is highly parallelizable and optimized for high-performance computing (HPC), making it scalable and faster for processing [39]. This format allows data to be stored in self-explanatory structures of the time-series images, where time-series Sentinel-2 images for each point can be processed individually with their corresponding attributes. Each HDF5 file contains datasets associated with LUCAS points from a single country in the form of a 3-D array of Sentinel-2 64×64 time-series patches, individually named as "$Lucas\_Point\_\{Point\_ID\}$" with their corresponding 119 attributes. Some examples of true color compositions of 64×64 Sentinel-2 patches belonging to different countries are shown in Fig. 6.

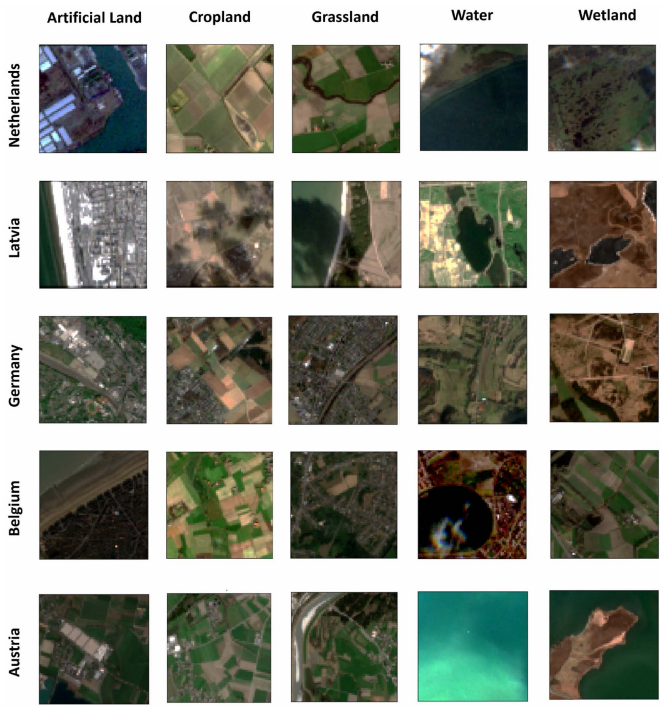


Fig. 6. True color composition of 64×64 Sentinel-2 patches belonging to the Sen4Map dataset, which shows the variability of the land-cover classes located in different countries.

IV. Land-Cover Mapping—Benchmarking

This section provides a comparison of the state-of-the-art classification algorithms most commonly used to generate land-cover maps, emphasizing the properties and challenges posed by the proposed Sen4Map dataset. In particular, we tested and compared:

1) the pixel-based Random Forest classifier [40], widely used for its robustness in the presence of noisy data [41];
2) the pixel-based transformer DL model [42], which is extremely effective in handling mapping tasks requiring time series of satellite data;
3) the standard spatial ViT model [43], typically utilized for image processing through its attention mechanism; and

TABLE I
DETAILED MAPPING OF LUCAS LABELS FOR BROAD LAND-COVER CLASSIFICATION

Target land-cover	LUCAS land-cover
Artificial Land	A- Artificial Land (Except A22)
Bareland and Lichens/Moss	A22 – Non built-up linear features F – Bareland and Lichens/Moss
Grassland	E - Grassland B50 – Fodder Crops
Cropland	B10 - Cereals B20 - Root Crops B30 - Non-Permanent Industrial Crop B40 – Dry Pulses, Vegetables and Flowers B70 – Permanent Fruits B80 – Other Permanent Crops BX1 BX2
Broadleaves	C10 – Broadleaved Woodland
Conifers	C20 – Coniferous Woodland C30 – Mixed Woodland CXX - LUCAS SU FT
Shrubland	D- Shrubland
Water	G – Water Areas
Wetlands	H - Wetlands

4) the Video Vision Transformer [44], capable of integrating both spatial and temporal information provided by time series of satellite data.

A. Experimental Setup

Thanks to the detailed description of the land-use and land-cover properties of the LUCAS samples, customized classification tasks can be defined. In the experimental setup we consider, we provide two examples: 1) the classification of the most widespread land-cover categories typically mapped at global scale [31], i.e., namely, "Water," "Broadleaves," "Shrubland," "Conifers," "Grass," "Crops," "Wetland," "Artificial Land," and "Bareland," and 2) a detailed crop classification task to map "Cereals," "Root Crops," "Non-permanent Industrial crop," "Dry pulses, vegetables, flowers," "Fodder crops," "Bareland," "Woodland and Shrubland," and "Grassland" according to the map legend presented in [45]. We describe how we generate these legends from the detailed LUCAS land-cover labels in Tables I and II, respectively. However, users have the flexibility to define alternative classification tasks based on their specific requirements.

The dataset used for training contains 70% of randomly selected points from each country. Similarly, from the exclusive remaining points per country, validation, and test sets were formed, each containing 15% of randomly selected exclusive data points from each country. In the case of the (1) land-cover classification task, we use 234 555 points for training, 50 284 points for validation, and 50 286 for testing. For the second crop classification task, we use 196 514 points for training, 48 394 points for validation, and 48 396 for testing, by considering only points having LUCAS labels as shown in Table II.

TABLE II
DETAILED MAPPING OF LUCAS LABELS FOR CROP CLASSIFICATION

Target land-cover	LUCAS land-cover
Cereals	B11 - Common wheat B12 - Durum wheat B13 - Barley B14 - Rye B15 - Oats B16 - Maize B17 - Rice B18 - Triticale B19 - Other cereals
Root Crops	B12 - Potatoes B22 - Sugar beet B23 - Other roots crops
Nonpermanent industrial crops	B34 - Cotton B35 - Other fibre and oleaginous crops B36 - Tobacco B37 - Other nonpermanent industrial crops
Dry pulses, vegetables and flowers	B31 - Sunflower B32 - Rape and turnip rape B33 - Soya B41 - Dry pulses B42 - Tomatoes B43 - Other fresh vegetables B44 - Floriculture and ornamental plants B45 - Strawberries
Fodder crops	B51 - Clovers B52 - Lucerne B53 - Other leguminous and mixtures for fodder B54 - Mixed cereals for fodder
Bareland	F10 - Rocks and Stone F20 - Sand F30 - Lichens and Moss F40 - Other Bare Soil
Woodland and Shrubland	B71-B77 - Fruit Trees B81- Olive Grooves B82 - Vineyards B83 - Nurseries B84 - Permanent industrial crops C10 - Broadleaved Woodland C20 - Coniferous Woodland C30 - Mixed Woodland D10 - Shrubland with sparse tree cover D20 - Shrubland without tree cover
Grassland	B55 - Temporary Grasslands E10 - Grassland with sparse tree/shrub cover E20 - Grassland without tree/shrub cover E30 - Spontaneously re-vegetated surfaces

The Random Forest classifier and the pixel-based transformer DL model were trained considering the center pixel of the 64×64 patch for both classification tasks. To harmonize the Sentinel-2 data available for different locations from the temporal point of view, a time series of 12 monthly composites was generated for each LUCAS point. To this end, we considered the widely used median composite approach described in [46], by masking out the clouds according to the information available in the "SCL" classification map associated with each Sentinel-2 image. However, we would like to remark that the proposed Sen4Map dataset provides the original time-series of Sentinel-2 images. Therefore, any approach can be used to harmonize the satellite data from the spatial and temporal viewpoint.

In contrast to the random forest and the pixel-based transformer DL model, the spatial ViT model was trained using 15×15 patches, which were cropped from the original 64×64 patches, and were associated with annual composites. To highlight the properties of the Sen4Map dataset, we choose a patch size of 15×15 as a proof of concept. The selection of a 15×15 patch was a reasonable setup for some specific methods such

TABLE III
MEAN ENTROPY (IN NAT) OBTAINED FOR DIFFERENT ML MODELS

Land-cover	Random Forest	Transformer (pixel-based)	Vision Transformer	Video Vision Transformer
Broad land-cover [31]	1.16	0.66	0.74	0.74
Crop land-cover [45]	0.83	0.39	0.72	0.65

TABLE IV
F-SCORE AND OVERALL ACCURACY OBTAINED ON THE SET OF LAND-COVER CLASSES TYPICALLY MAPPED [31]

Classes	Random Forest		Transformer (pixel-based)		ViT		VViT	
	Center pixel	3 x 3 window	Center pixel	3 x 3 window	Center pixel	3 x 3 window	Center pixel	3 x 3 window
Artificial land	0.49	0.72	0.57	0.78	0.53	0.77	0.59	0.79
Bareland	0.2	0.38	0.24	0.47	0.2	0.43	0.25	0.5
Broadleaves	0.69	0.86	0.73	0.9	0.69	0.84	0.75	0.87
Conifers	0.76	0.88	0.80	0.91	0.78	0.87	0.81	0.9
Cropland	0.8	0.89	0.83	0.92	0.78	0.87	0.83	0.9
Grassland	0.69	0.84	0.73	0.89	0.68	0.84	0.73	0.87
Shrubland	0.29	0.51	0.42	0.73	0.31	0.54	0.43	0.65
Water	0.61	0.74	0.63	0.79	0.6	0.75	0.65	0.79
Wetlands	0.6	0.75	0.67	0.87	0.61	0.76	0.7	0.83
W.A. F-score	0.67	0.81	0.72	0.87	0.67	0.81	0.72	0.85
Overall Accuracy	0.68	0.83	0.73	0.88	0.68	0.82	0.73	0.85

TABLE V
F-SCORE AND OVERALL ACCURACY OBTAINED ON THE CROP CLASSIFICATION LEGEND PROPOSED IN [45]

Classes	Random Forest		Transformer (pixel-based)		ViT		VViT	
	Center pixel	3 x 3 window	Center pixel	3 x 3 window	Center pixel	3 x 3 window	Center pixel	3 x 3 window
Cereals	0.78	0.88	0.83	0.93	0.66	0.78	0.79	0.86
Root crops	0.56	0.69	0.69	0.88	0.24	0.44	0.64	0.75
Industrial crops	0.64	0.76	0.72	0.88	0.42	0.58	0.67	0.78
Dry pulses,flowers	0.07	0.1	0.31	0.68	0.04	0.08	0.18	0.35
Fodder crops	0.01	0.01	0.31	0.64	0.19	0.34	0.37	0.54
Bareland	0.33	0.47	0.40	0.64	0.21	0.37	0.32	0.47
Wood/Shrubland	0.88	0.96	0.90	0.97	0.85	0.93	0.88	0.95
Grassland	0.71	0.88	0.76	0.93	0.66	0.84	0.71	0.87
W.A. F-score	0.77	0.88	0.82	0.93	0.69	0.82	0.76	0.86
Overall Accuracy	0.79	0.89	0.82	0.93	0.70	0.83	0.77	0.87

as patch-based transformer models that later allowed us to emphasize the attention to segments of smaller 3×3 patches, which later helped with the inference and comparison with other models.

Similar to the monthly composites, the annual composite was calculated by computing the median value of all the cloud-free Sentinel-2 images associated with the LUCAS location according to the cloud information provided by the "SCL" classification map. Finally, to leverage both the spatial and temporal information available in the Sen4Map dataset, the VViT model was trained on 15×15 patches (cropped from 64×64 patches) associated with the time series of 12 monthly composites. The need for harmonization of the data from a temporal viewpoint is based on the architectural limitation of the VViT model. Cross-entropy [47] is used to optimize the transformer-based models. A description of important hyperparameters used for training ML models is provided in the Appendix. In addition, to assess uncertainties in the performance of different classification algorithms, Entropy [48] is reported as shown in Table III. Higher value indicates higher uncertainty.

B. Experimental Results

Tables IV and V report the experimental results obtained with the different state-of-the-art classification methods on the land-cover legend proposed in [31], and the crop classification legend proposed in [45]. Differently from Venter et al. [31], the "Tree" class is separated into "Broadleaves" and "Conifers." The obtained predictions are evaluated by comparing the land-cover information with the classification

Fig. 7. Examples of conifer points. The 3×3 window allows the correct evaluation of the land-cover classification result with the real land-cover present in the scene.

results associated with exactly the LUCAS point location and a 3×3 window centered in the LUCAS point location. In the second case, if at least one of the pixels is associated with the LUCAS land-cover label, the point is considered correctly classified. This condition allows us to mitigate possible GPS errors and handle classes for which the LUCAS observation is related to a land-cover within an observation window in an area defined by a radius of 20 m around the point. Fig. 7 shows the importance of evaluating the classification results considering the 3×3 pixel window by reporting an example of two LUCAS points associated with conifer land-cover. This can be validated by comparing the overall performance of all classification methods using the center pixel, as superior results are achieved with a 3×3 window. Furthermore, all classification methods exhibited similar performance for their respective classes.

By focusing on the land-cover classification task, it is evident that the most challenging classes are Bareland and Shrubland. The classification of Shrublands poses difficulties due to the spatial disparity between the satellite's resolution and the ground-level land-cover information. To address this issue, many approaches documented in the literature that utilized the LUCAS database to train their models often employed spatial filtering techniques to refine both the training and validation datasets, thereby improving accuracy. Despite this, transitional classes such as Bareland and Shrubland, which exhibit seasonal and spatial variability, are highly susceptible to misclassification. This phenomenon is also noticeable in classes such as Dry pulses, flowers, and Fodder crops within the crop classification task.

V. Discussion

One of the most significant challenges posed by the proposed dataset Sen4Map pertains to the spatial disparity between the resolution of satellite imagery and the scale of ground-based land-cover information. Unlike existing benchmark datasets, which are generated through photointerpretation of satellite images for land-cover mapping, Sen4Map relies on in situ data. The implementation of a 3×3 window, described in Section IV-B, has been demonstrated to mitigate the disparities between satellite spatial resolution and ground-based in situ

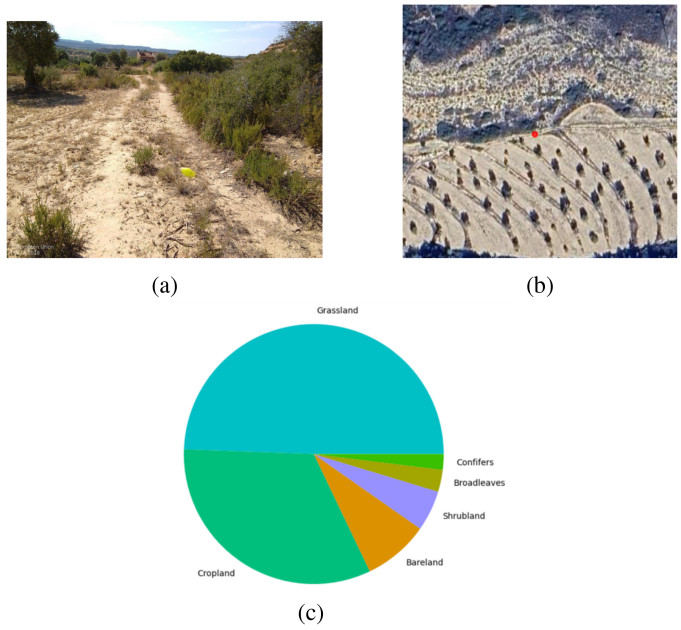
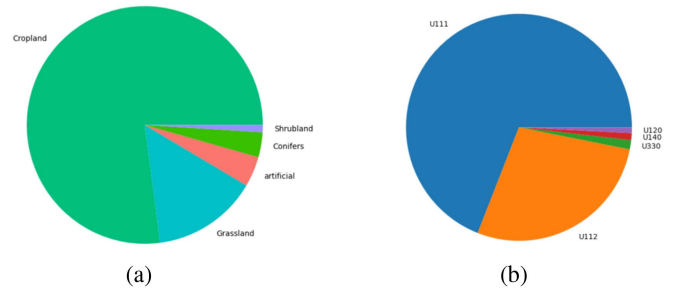


Fig. 8. Land-cover classification task: An example of misclassification from the Sen4Map dataset is depicted. (a) Shows an in-situ street-level image from LUCAS, (b) displays a very high resolution Google Map image, and (c) illustrates the number of misclassified points categorized as land-use "U111," which belong to the class "Artificial land".

observations. However, unlike existing methodologies [33], we did not apply any data cleaning procedures. This decision aligns with the dataset's objective of advancing mapping capabilities using Sentinel-2 data, particularly on challenging land-cover samples.

In addition, the Sen4Map dataset offers the possibility to analyze classification outcomes based on their comprehensive land-use and land-cover description. For instance, Fig. 8 shows a LUCAS location highlighted in red in Fig. 8(b), associated with the land-cover label **Lc1** ("A21"—Non-built-up area features), belonging to the class "Artificial land," but misclassified as "Cropland." In the LUCAS in situ street-level picture shown in Fig. 8(a), one can observe that although the LUCAS location is correctly associated with a road (belonging to the "Artificial land" category), it is actually sealed land within an agricultural landscape. This property can also be identified by its description. The point is associated with a land-use label **Lu1** equal to "U312" (road transport) and **Lu2** equal to "U111" (agriculture excluding fallow land and kitchen areas). This land-use description is highly informative for understanding the subset of misclassified samples categorized as "Artificial land" and misclassified labeled as "Cropland." Fig. 8(c) displays the set of misclassified samples with land-use code "U111," belonging to the "Artificial land" class. It is evident that these points are primarily misclassified as "Grassland" and "Cropland" when only considering the center pixel.

To understand the challenges associated with classifying land cover types like "Bareland," Fig. 9(a) presents the collective misclassification occurrences linked to the "Bareland" class, specifically concerning the land-cover label **Lc1** identified as



(a) (b)

Fig. 9. Land-cover classification analysis: examining misclassified samples in the "Bareland" class. (a) Misclassifications within the "Bareland" class associated to the land-cover label **Lc1** "F40". (b) Land-use associated with misclassified samples labeled as "F40," incorrectly identified as "Cropland."

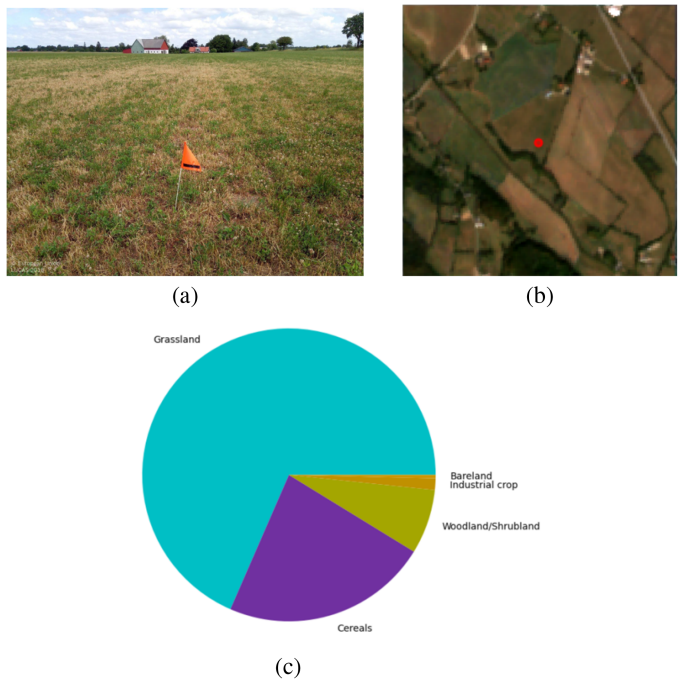
(a) (b)

Fig. 10. Example from Sen4Map dataset from Bareland class having land-cover "F40" and land-use "U111," misclassified as Cropland. (a) LUCAS in situ street-level image. (b) Yearly composite of Sentinel-2 image.

"F40" (Other Bare Soil). A considerable portion of the misclassified samples labeled as "F40" are erroneously identified as "Cropland." For instance, Fig. 10 shows a location associated with LUCAS land-cover **Lc1** "F40" and land-use **Lu1** "U111" that was misclassified as "Cropland."

To analyze further, the associated land-use from those samples that were misclassified as "Cropland" is plotted in Fig. 9(b). It can be seen that mainly "U111" and "U112" are responsible for these misclassifications since they clarify that this specific land-cover is associated with "Agricultural production" and "Fallow agricultural land," respectively. Indeed, in the LUCAS classification scheme, the land-cover label "F40" links to land-use "U111" as agriculture bare land, tilled and/or prepared for seeding, and "U112" as fallow land with crop residues, becoming the main reason for such misclassifications.

In a comparable scenario involving crop classification, particularly for the challenging class "Fodder crop," a majority of misclassifications from the "Fodder crop" were identified as "Grassland," as depicted in Fig. 11(c), when considering predictions based on a center pixel. This observation is further supported by Fig. 11(a) and (b), where a sample with the land-cover label **Lc1** "B53" (other legumes and mixtures for fodder) exhibits a notable semblance to temporary grasslands when scrutinized through a satellite-based viewpoint. Since these crops are highly prone to seasonal variations, the temporal component is crucial in determining accurate classification predictions. This can be validated from Table V, where it is evident



(a) (b) (c)

Fig. 11. Crop classification task: detailed analysis of misclassified samples from class "Fodder crops." (a) LUCAS in situ street-level image and (b) yearly composite of Sentinel-2 image. (c) All misclassifications from class "Fodder crops."

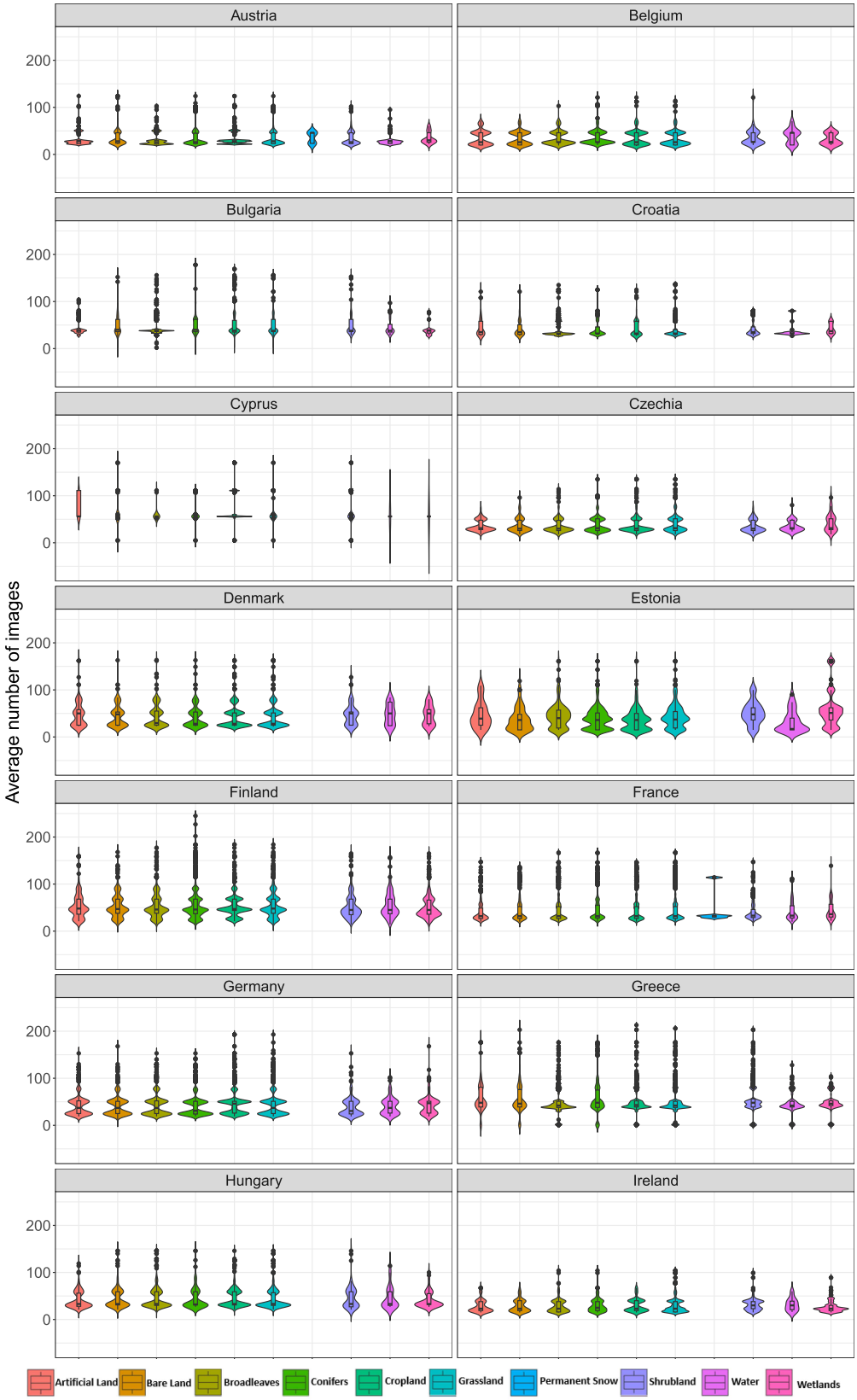
that in cases of pixel-based transformers and video-vision transformers, where the attention mechanism is highly dependent on the temporal dimension, results in better predictions. In addition, a 3×3 window further improves the classification results. Further observations within a 3×3 window indicate that the chances of correct classification increase by at least 15%.

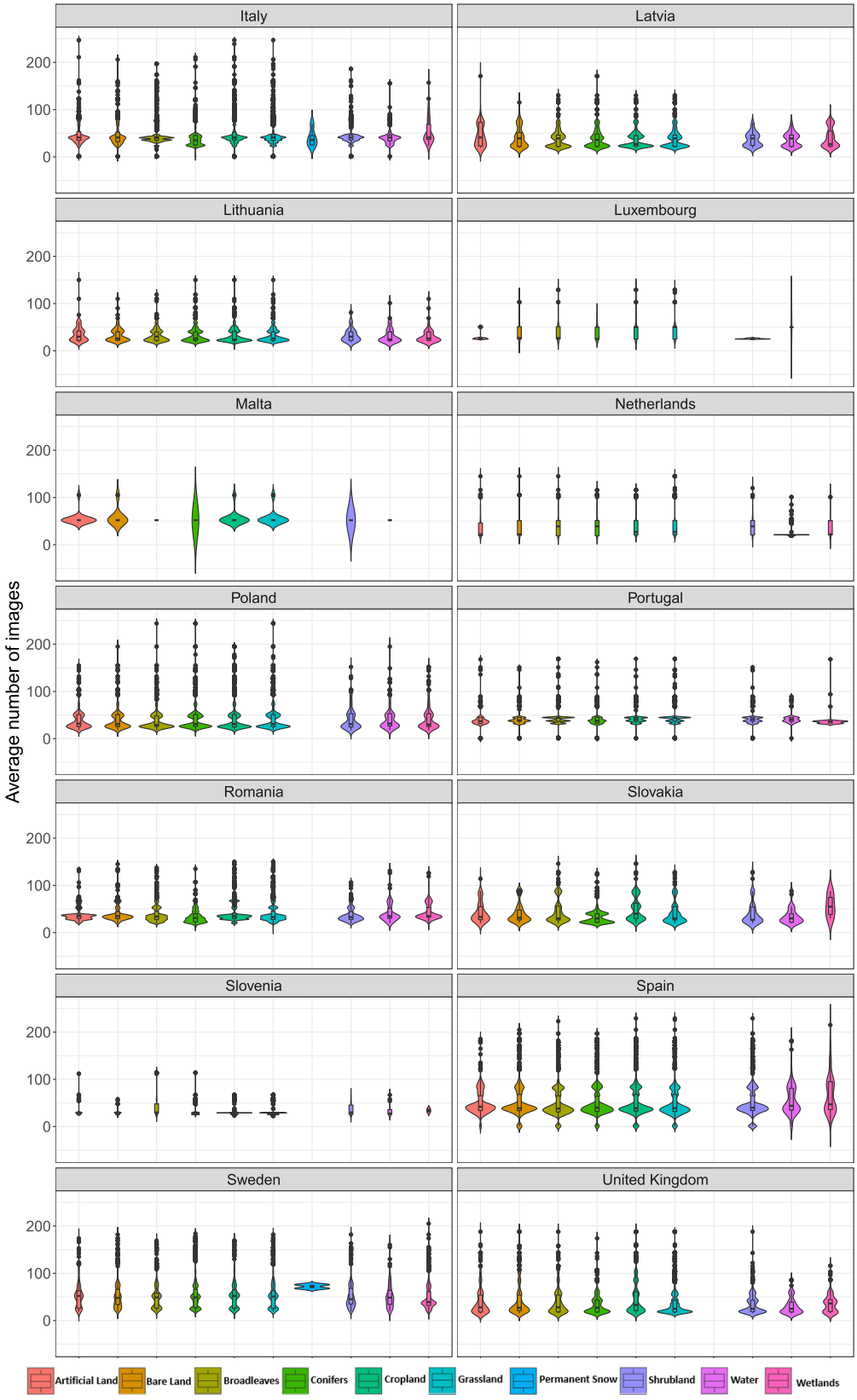
VI. Conclusion

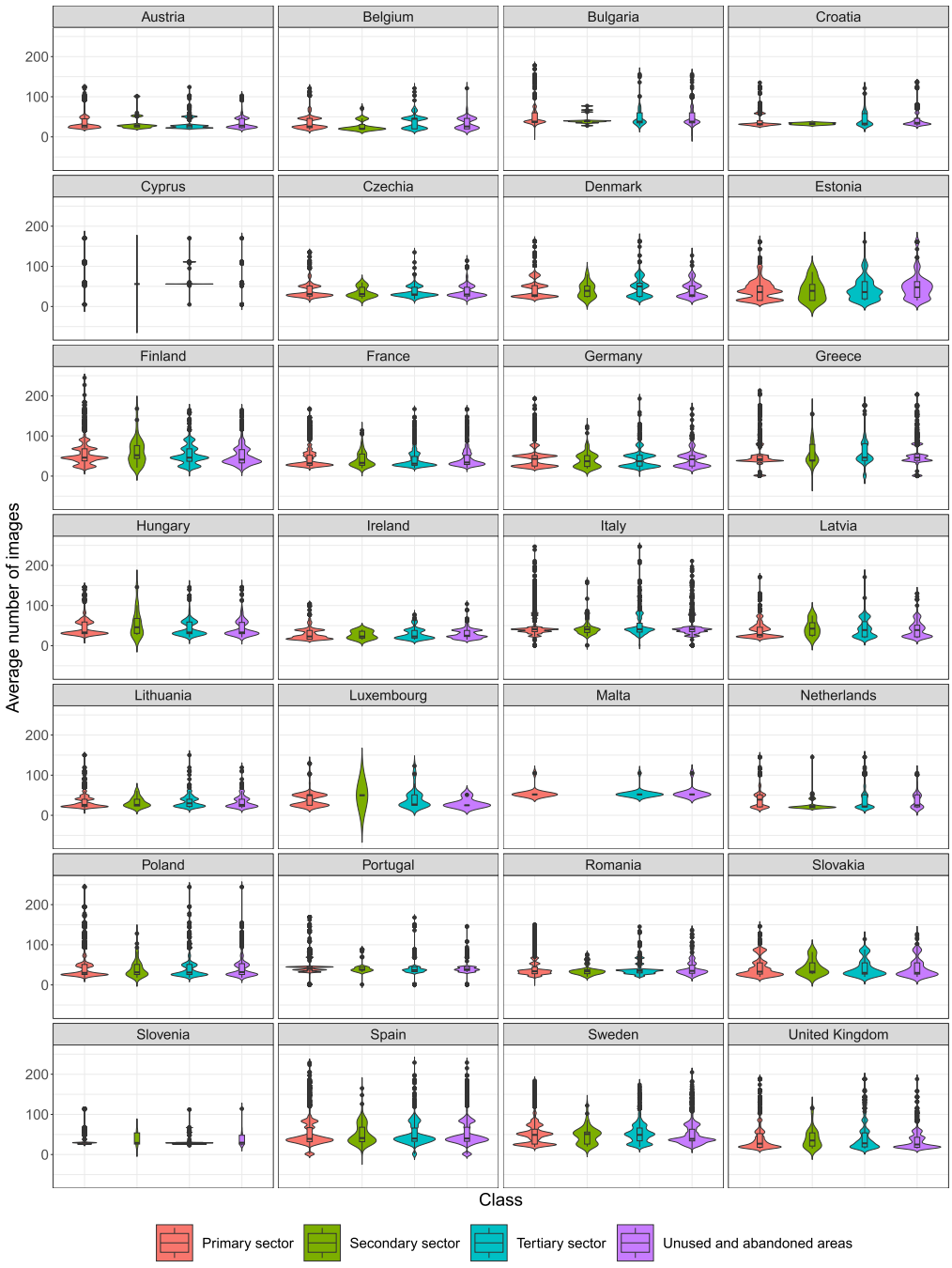
To enhance the capability of generating land-cover maps using Sentinel-2 data, this article presents Sen4Map, a large-scale benchmark dataset. It comprises nonoverlapping 64×64 patches extracted from Sentinel-2 time series images, covering over 335 125 geotagged locations across the European Union. These geotagged locations are associated with detailed land-cover and land-use information, gathered by experts as part of the LUCAS survey in 2018. The proposed dataset is highly customizable, offering flexibility across different land-cover and land-use categories. It supports both spatial and temporal classification approaches and is scale independent. The dataset is available in countrywise, highly parallelizable HDF5 files, showcasing high modularity in terms of scale, and land-cover and land-use information. This article demonstrates the dataset's application through various classification approaches using state-of-the-art models for generating land-cover maps. Furthermore, it highlights the challenges posed by the schematic gap between satellite and in situ data, offering opportunities for developing techniques and methods to advance current land-cover mapping efforts with Sentinel-2.

APPENDIX

Countrywise statistical distribution of the number of Sentinel-2 images associated with each land-cover and land-use class considering the first level of the hierarchy in Sen4Map benchmark dataset, depicted as a violin plot.




Italy
Latvia
Lithuania
Luxembourg
Malta
Netherlands
Poland
Portugal
Romania
Slovakia
Slovenia
Spain
Sweden
United Kingdom
Average number of images
0
100
200
Artificial Land
Bare Land
Broadleaves
Conifers
Cropland
Grassland
Permanent Snow
Shrubland
Water
Wetlands


Austria
Belgium
Bulgaria
Croatia
Cyprus
Czechia
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Ireland
Italy
Latvia
Lithuania
Luxembourg
Malta
Netherlands
Poland
Portugal
Romania
Slovakia
Slovenia
Spain
Sweden
United Kingdom
Average number of images
0
100
200
Class
Primary sector
Secondary sector
Tertiary sector
Unused and abandoned areas

The list of some of the hyperparameters used for training ML models is shown in tabular form.

Hyperparameters	Random Forest
Number of trees	400
criterion	gini
Max depth	Until pure
Minimum number of splits	2
Maximum number of features	sqrt

Hyperparameters	Transformer (pixel-based)
Embedding Dimension	128
Hidden Dimension	256
Number of heads	4
Number of layers	3
Number of patches	12
Learning rate	3e-4
Dropout	0.2

Hyperparameters	Vision Transformer
Embedding Dimension	256
Hidden Dimension	512
Number of heads	8
Number of layers	6
Image Patch size	3
Number of patches	25
Learning rate	3e-4
Dropout	0.2

Hyperparameters	Video Vision Transformer
Image size	15
Frames	12
Image Patch size	3
Frame Patch size	3
Embedding Dimension	256
Hidden Dimension	512
Number of heads	8
Spatial Depth	3
Temporal Depth	3
Learning rate	3e-5

ACKNOWLEDGMENT

The authors would like to thank the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

REFERENCES

[1] C. Persello et al., "Deep learning and earth observation to support the sustainable development goals: Current approaches, open challenges, and future opportunities," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 172–200, Jun. 2022.

[2] C. Paris, L. Bruzzone, and D. Fernández-Prieto, "A novel approach to the unsupervised update of land-cover maps by classification of time series of multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4259–4277, Jul. 2019.

[3] M. Russwurm, N. Courty, R. Emonet, S. Lefèvre, D. Tuia, and R. Tavenard, "End-to-end learned early classification of time series for in-season crop type mapping," *ISPRS J. Photogrammetry Remote Sens.*, vol. 196, pp. 445–456, 2023.

[4] Copernicus, "Sentinel-2 mission overview," 2024, Accessed: Jun. 12, 2024. [Online]. Available: https://sentiwiki.copernicus.eu/web/s2-mission

[5] M. Rußwurm, C. Pelletier, M. Zollner, S. Lefèvre, and M. Körner, "Breizhcrops: A tome series dataset for crop type mapping," in *XXIV ISPRS Congr., Commission II*, (International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives), N. Paparoditis, C. Mallet, F. Lafarge, F. Remondino, I. Toschi, and T. Fuse, Eds., Aug. 2020, pp. 1545–1551.

[6] V. Akbari, S. Solberg, and S. Puliti, "Multitemporal Sentinel-1 and Sentinel-2 images for characterization and discrimination of young forest stands under regeneration in Norway," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 5049–5063, 2021.

[7] J. Parajuli, R. Fernandez-Beltran, J. Kang, and F. Pla, "Attentional dense convolutional neural network for water body extraction from Sentinel-2 images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 6804–6816, 2022.

[8] C. F. Brown et al., "Dynamic world, near real-time global 10 m land use land cover mapping," *Sci. Data*, vol. 9, no. 1, pp. 1–17, 2022.

[9] R. C. Sharma, "Countrywide mapping of plant ecological communities with 101 legends including land cover types for the first time at 10m resolution through convolutional learning of satellite images," *Appl. Sci.*, vol. 12, no. 14, 2022, Art. no. 7125. [Online]. Available: https://www.mdpi.com/2076-3417/12/14/7125

[10] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "SEN12MS–a curated dataset of georeferenced multi-spectral Sentinel-1/2 imagery for deep learning and data fusion," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. IV-2/W7, pp. 153–160, 2019. [Online]. Available: https://isprs-annals.copernicus.org/articles/IV-2-W7/153/2019/

[11] P. Ebel, Y. Xu, M. Schmitt, and X. X. Zhu, "SEN12MS-CR-TS: A remote-sensing data set for multimodal multitemporal cloud removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5222414.

[12] C. Aybar et al., "CloudSEN12, A global dataset for semantic understanding of cloud and cloud shadow in Sentinel-2," *Sci. Data*, vol. 9, no. 1, Dec. 2022, Art. no. 782.

[13] J. Cornebise, I. Orsolic, and F. Kalaitzis, "Open high-resolution satellite imagery: The worldstrat dataset—With application to super-resolution," in *Proc. 36th Conf. Neural Inf. Process. Syst. Datasets Benchmarks Track*, 2022, pp. 25979–25991.

[14] Y. Cong et al., "Satmae: Pre-training transformers for temporal and multi-spectral satellite imagery," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 197–211, 2022.

[15] O. Mañas, A. Lacoste, X. GiróI Nieto, D. Vázquez, and P. Rodríguez López, "Seasonal contrast: Unsupervised pre-training from uncurated remote sensing data," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 9394–9403.

[16] Y. Wang, N. A. A. Braham, Z. Xiong, C. Liu, C. M. Albrecht, and X. X. Zhu, "SSL4EO-S12: A large-scale multimodal, multitemporal dataset for self-supervised learning in earth observation [software and data sets]," *IEEE Geosci. Remote Sens. Mag.*, vol. 11, no. 3, pp. 98–106, Sep. 2023.

[17] F. Bastani, P. Wolters, R. Gupta, J. Ferdinando, and A. Kembhavi, "Satlaspretrain: A large-scale dataset for remote sensing image understanding," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2022, pp. 16772–16782.

[18] A. Francis and M. Czerkawski, "Major tom: Expandable datasets for earth observation," 2024, *arXiv:2402.12095*.

[19] J. Munoz-Mari et al., "Hyperlabelme : A web platform for benchmarking remote-sensing image classifiers," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 79–85, Dec. 2017.

[20] J. Jakubik et al., "Foundation models for generalist geospatial artificial intelligence," 2023, *arxiv:2310.18660*.
[21] G. Sumbul, M. Charfuelan, B. Demir, and V. Markl, "Bigearthnet: A large-scale benchmark archive for remote sensing image understanding," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 5901–5904.
[22] 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.
[23] R. Van De Kerchove et al., "Esa worldcover: Global land cover mapping at 10 m resolution for 2020 based on Sentinel-1 and 2 data," in *AGU Fall Meeting Abstr.*, vol. 2021, 2021, pp. GC45I–0915.
[24] K. Karra, C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby, "Global land use/land cover with Sentinel 2 and deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 4704–4707.
[25] M. Probeck et al., "CLC backbone: Set the scene in copernicus for the coming decade," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 2076–2079.
[26] C. Pelletier, G. I. Webb, and F. Petitjean, "Temporal convolutional neural network for the classification of satellite image time series," *Remote Sens.*, vol. 11, no. 5, 2019, Art. no. 523.
[27] C. Paris and L. Bruzzone, "Automatic extraction of weak labeled samples from existing thematic products for training convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 5722–5725.
[28] Copernicus, "Sentinel-2A, processed by ESA, MSI level-2A boa reflectance product. Collection 1, European Space Agency," 2024, Accessed: Jun. 12, 2024. [Online]. Available: https://doi.org/10.5270/S2_-znk9xsj
[29] R. D'Andrimont et al., "Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European union," *Sci. Data*, vol. 7, no. 1, 2020, Art. no. 352.
[30] Y. Gao, L. Liu, X. Zhang, X. Chen, J. Mi, and S. Xie, "Consistency analysis and accuracy assessment of three global 30-m land-cover products over the European union using the LUCAS dataset," *Remote Sens.*, vol. 12, no. 21, 2020, Art. no. 3479.
[31] Z. S. Venter, D. N. Barton, T. Chakraborty, T. Simensen, and G. Singh, "Global 10m land use land cover datasets: A comparison of dynamic world, world cover and ESRI land cover," *Remote Sens.*, vol. 14, no. 16, 2022, Art. no. 4101.
[32] M. Weigand, J. Staab, M. Wurm, and H. Taubenböck, "Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 88, 2020, Art. no. 102065.
[33] Z. S. Venter and M. A. Sydenham, "Continental-scale land cover mapping at 10m resolution over Europe (ELC10)," *Remote Sens.*, vol. 13, no. 12, 2021, Art. no. 2301.
[34] S. M. Mirmazloumi et al., "ELULC-10, a 10m European land use and land cover map using Sentinel and landsat data in google earth engine," *Remote Sens.*, vol. 14, no. 13, 2022, Art. no. 3041.
[35] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425717302900
[36] C. Ji et al., "Time series classification with random temporal features," *J. King Saud Univ.— Comput. Inf. Sci.*, vol. 35, no. 9, 2023, Art. no. 101783. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1319157823003373
[37] Z. Liu, X. Li, and Y. Cui, "Time series model interpretation via temporal feature sampling," in *Web and Big Data*, X. Song, R. Feng, Y. Chen, J. Li, and G. Min, Eds. Singapore: Springer Nature Singapore, 2024, pp. 346–360.
[38] M. Herold, R. Hubald, and A. Di Gregorio, "Translating and evaluating land cover legends using the UN land cover classification system LCCS," *GOFC-GOLD Rep.*, vol. 43, pp. 1482–1483, 2009.
[39] The HDF Group, "Hierarchical data format, version 5," 1997. Accessed: Jun. 12, 2024. [Online]. Available: http://www.hdfgroup.org/HDF5/. [Online]. Available: http://www.hdfgroup.org/HDF5/
[40] C. Pelletier, S. Valero, J. Inglada, N. Champion, C. Marais Sicre, and G. Dedieu, "Effect of training class label noise on classification performances for land cover mapping with satellite image time series," *Remote Sens.*, vol. 9, no. 2, 2017, Art. no. 173.
[41] R. C. Sharma, "An ultra-resolution features extraction suite for community-level vegetation differentiation and mapping at a sub-meter resolution," *Remote Sens.*, vol. 14, no. 13, 2022, Art. no. 3145. [Online]. Available: https://www.mdpi.com/2072-4292/14/13/3145
[42] A. Vaswani et al., "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, (NIPS17), Red Hook, NY, USA: Curran Associates Inc., 2017, p. 6000–6010.
[43] A. Dosovitskiy et al., "An image is worth 16 x 16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Representations*, 2021. [Online]. Available: https://openreview.net/forum?id=YicbFdNTTy
[44] A. Arnab, M. Dehghani, G. Heigold, C. Sun, M. Lučić, and C. Schmid, "Vivit: A video vision transformer," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 6836–6846.
[45] R. d'Andrimont, A. Verhegghen, G. Lemoine, P. Kempeneers, M. Meroni, and M. Van Der Velde, "From parcel to continental scale–a first European crop type map based on sentinel-1 and LUCAS copernicus in-situ observations," *Remote Sens. Environ.*, vol. 266, 2021, Art. no. 112708. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425721004284
[46] G. Weikmann, C. Paris, and L. Bruzzone, "TimeSen2Crop: A million labeled samples dataset of Sentinel 2 image time series for crop-type classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4699–4708, 2021.
[47] R. Y. Rubinstein and D. P. Kroese, *The Cross Entropy Method: A. Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2004.
[48] A. Namdari and Z. S. Li, "A review of entropy measures for uncertainty quantification of stochastic processes," *Adv. Mech. Eng.*, vol. 11, no. 6, 2019, Art. no. 1687814019857350. [Online]. Available: https://doi.org/10.1177/1687814019857350

Surbhi Sharma (Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from Amity University, Noida, India, in 2015, and the M.Sc. degree in geoinformation science and earth observation with a specialization in geoinformatics jointly from the Indian Institute of Remote Sensing, Indian Space and Research Organization, Bangalore, India, and the Faculty of Geo-information Science and Earth Observation, University of Twente, Enschede, The Netherlands, in 2018. She is currently working toward the Ph.D. degree in computational engineering with the University of Iceland, Reykjavik, Iceland.

She is a Member of the "AI and ML for Remote Sensing" Simulation and Data Lab, Jülich Supercomputing Centre, Germany. Her research interest lies in scalable machine learning and deep learning methods for remote sensing applications, with a particular focus on advanced deep transfer learning methods using modern high-performance computing systems.

Rocco Sedona (Member, IEEE) received the B.Sc. and M.Sc. degrees in information engineering from the University of Trento, Trento, Italy, in 2016 and 2019, respectively, and the Ph.D. degree in computational engineering from the University of Iceland, Reykjavik, Iceland, in 2023.

He is a member of the "AI and ML for Remote Sensing" Simulation and Data Lab, JSC, Germany. His research interests primarily lie in the field of deep learning and its application to remote sensing data. He has extensively utilized optical satellite data acquired by Landsat (NASA) and Sentinel (ESA) missions toward near real-time land-cover classification. In addition, he specializes in distributed deep learning on high-performance computing systems, an area of study that he has been actively engaged in since 2019.

Morris Riedel (Member, IEEE) received the Ph.D. degree in computer engineering from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, and worked in data-intensive parallel and distributed systems since 2004.

He is currently a Full Professor of high-performance computing with an emphasis on parallel and scalable machine learning with the School of Natural Sciences and Engineering, University of Iceland, Iceland. Since 2004, he held various positions at the Juelich Supercomputing Centre of Forschungszentrum Juelich, Germany. In addition, he is the Head of the joint High Productivity Data Processing Research Group between the Juelich Supercomputing Centre and the University of Iceland. Since 2020, hehas been the EuroHPC Joint Undertaking governing board member for Iceland. His research interests include high-performance computing, remote sensing applications, medicine and health applications, pattern recognition, image processing, and data sciences, and he has authored extensively in those fields. His online YouTube and university lectures include high-performance computing—advanced scientific computing, cloud computing and big data—parallel and scalable machine and deep learning, as well as statistical data mining. In addition, he has performed numerous hands-on training events in parallel and scalable machine and deep learning techniques on cutting-edge HPC systems.

Gabriele Cavallaro (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in telecommunications engineering from the University of Trento, Trento, Italy, in 2011 and 2013, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Iceland, Iceland, in 2016.

From 2016 to 2021, he has been the Deputy Head of the "High Productivity Data Processing" (HPDP) Research Group, Jülich Supercomputing Centre, Forschungszentrum Jülich, Germany. Since 2022, he has been the Head of the "AI and ML for Remote Sensing" Simulation and Data Lab at the JSC and an Adjunct Associate Professor with the School of Natural Sciences and Engineering, University of Iceland, Iceland. From 2020 to 2023, he held the position of the Chair for the High-Performance and Disruptive Computing in Remote Sensing Working Group under the IEEE GRSS Earth Science Informatics Technical Committee. In 2023, he took on the role of Co-chair for the ESI TC. Concurrently, he was a Visiting Professor with the Φ-lab within the European Space Agency (ESA), where he contributes to the Quantum Computing for Earth Observation initiative. In addition, he is an Associate Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING since October 2022. His research interests cover remote sensing data processing with parallel machine learning algorithms that scale on distributed computing systems and innovative computing technologies.

Dr. Cavallaro was the recipient of the IEEE GRSS Third Prize in the Student Paper Competition of the IEEE International Geoscience and Remote Sensing Symposium 2015 (Milan—Italy).

Claudia Paris (Senior Member, IEEE) received the B.S. and M.S. (*summa cum laude*) degrees in telecommunication engineering and the Ph.D. degree in information and communication technology from the University of Trento, Trento, Italy, in 2010, 2012, 2016, respectively. She accomplished the Honors Master Program in research within the master's degree in telecommunication engineering in 2012.

Since 2021, she has been an Assistant Professor with the Faculty of Geoinformation and Earth Observation Sciences, University of Twente, Enschede, The Netherlands. Her research encompasses image and signal processing, machine learning and deep learning, specifically applied to remote sensing image analysis. She focuses on designing innovative and automated workflows for the analysis and classification of large-scale Earth observation data for various applications (e.g., forest/agricultural mapping and monitoring). To this end, she works with multisource remote sensing data (e.g., LiDAR, hyperspectral, multispectral, and high-resolution optical images), multitemporal image analysis, domain adaptation methods, development of ad hoc deep learning solutions, and biophysical parameter estimation. She conducts research on these topics within the frameworks of national and international projects. Since 2023, she has been an Associate Editor for the IEEE GEOSCIENCE OF REMOTE SENSING LETTERS.

Dr. Paris won the very prestigious Symposium Prize Paper Award at the 2016 International Symposium on Geoscience and Remote Sensing, Beijing, China, in 2016 and at the 2017 International Symposium on Geoscience and Remote Sensing, Fort Worth, TX, USA, in 2017. She also received the IEEE Geoscience and Remote Sensing Society 2022 Letters Prize Paper Award (exceptional paper in terms of content and impact on the GRS-Society). Since 2020, she has been a member of the program and scientific committee of several international conferences and workshops.