TimeSen2Crop: A Million Labeled Samples Dataset of Sentinel 2 Image Time Series for Crop-Type Classification

Giulio Weikmann, Claudia Paris, *Member, IEEE*, and Lorenzo Bruzzone, *Fellow, IEEE*

***Abstract*—This article presents TimeSen2Crop, a pixel-based dataset made up of more than 1 million samples of Sentinel 2 time series (TSs) associated to 16 crop types. This dataset, publicly available, aims to contribute to the worldwide research related to the supervised classification of TSs of Sentinel 2 data for crop type mapping. TimeSen2Crop includes atmospherically corrected images and reports the snow, shadows, and clouds information per labeled unit. The provided TSs represent an agronomic year (i.e., period from one year's harvest to the next one for agricultural commodity) ranging from September 2017 to August 2018. To generate the dataset, the publicly available Austrian crop type map based on farmer's declarations has been considered. To ensure the selection of reliable labeled units from the map (i.e., pure pixels correctly associated to their labels), an automatic procedure for the extraction of the training set based on a multitemporal deep learning model has been defined. TimeSen2Crop also includes a TS of Sentinel 2 images acquired in the following agronomic year (i.e., from September 2018 to August 2019). These data are provided with the aim of attract more research activities for solving a typical challenge of the crop type mapping task: adapting multitemporal deep learning models to different year (domain adaptation). The design of the dataset is described along with a benchmark comparison of deep learning models for crop type mapping.**

***Index Terms*—Benchmark, crop type mapping, multispectral images, multitemporal deep learning, Sentinel-2 dataset, time series (TSs), TimeSen2Crop.**

I. Introduction

SENTINEL 2 satellite constellation acquires multispectral images with high spatial and temporal resolutions. Having 13 spectral bands, which has a spatial resolution ranging from 10 to 60 m, and a revisit time of 5 days (depending on the latitude), Sentinel 2 generates dense time series (TSs) of images at global scale. Differently from similar optical earth observation missions (e.g., SPOT and Landsat), it acquires three bands in the red-edge spectral range, which provide key information for vegetation analysis. Given its temporal, spatial, and spectral resolutions, Sentinel 2 enables seasonal trend analysis, which are extremely useful for crop type mapping. However, the lack of large-scale training datasets hampers the possibility of developing and testing advanced methods for agricultural monitoring, such as those based on deep learning models. Indeed, to successfully train deep learning architectures, a large amounts of high-quality training data are required.

Recently, the RS community devoted a large effort to release large-scale benchmark datasets. Most of them focus on the modeling different application scenarios with single time RS images neglecting the temporal component. In [1], Helber *et al.* presents the EuroSAT dataset made up of 27 000 labeled and geo-referenced Sentinel 2 satellite image patches (i.e., 64 $\times$ 64 pixels). Although the classification scheme is made up of ten different classes, including land covers having peculiar temporal patterns (i.e., annual crops, permanent crops), the dataset is based on single-time images. In [2], the DeepGlobe2018 benchmark dataset provides 6867.6 million of labeled samples related to very high resolution satellite images (i.e., 50 cm). The dataset is made up of 10 146 RGB images of size 20448 $\times$ 20448 pixels associated to manual annotations based on pixelwise segmentation masks. No spectral or temporal information is provided.

Recently, by taking advantage from the open and free images acquired by the Sentinel satellites, a large-scale training dataset made up of geo-referenced Sentinel 1 SAR images and Sentinel 2 multispectral images has been released [3]. The dataset aims to facilitate the development of deep learning methods capable to extract and exploit the complementary information provided by multisensor RS data. Although the data employed have been acquired between December 2016 and November 2017, the temporal information was used only to split the data into four seasons, namely, winter (1 December 2016 to 28 February 2017), spring (1 March 2017 to 30 May 2017), summer (1 June 2017 to 31 August 2017), and fall (1 September 2017 to 30 November 2017). Similarly in [4], Sumbul *et al.* propose a large-scale training set for testing and developing novel methods in the context of the multilabel image classification task by considering Sentinel 2 images acquired between June 2017 and May 2018. Also in this case, the temporal information is used only to represent the land cover classes in different seasons.

In the context of the TS analysis, few benchmark datasets are available. The 2019 MediaEval Benchmark dataset was defined for monitoring flooding events [5]. The sequences of satellite

Manuscript received March 3, 2021; revised April 6, 2021; accepted April 13, 2021. Date of publication April 19, 2021; date of current version May 24, 2021. This work was supported in part by the European Union's Horizon 2020 research and innovation programme under Grant 825258. *(Corresponding author: Lorenzo Bruzzone.)*

The authors are with the Department of Information Engineering and Computer Science, University of Trento, Trento 38123, Italy (e-mail: giulio.weikmann@unitn.it; claudia.paris@unitn.it; lorenzo.bruzzone@ing.unitn.it).

Digital Object Identifier 10.1109/JSTARS.2021.3073965

images provided are focused on a fixed length of time around a flood event. The benchmark aims to test the capability of existing methods to create binary maps depicting the flooding event ongoing in an urban area. In [6], the Onera Satellite Change Detection dataset is proposed for urban change detection on Sentinel 2 images. The dataset includes 24 regions of approximately 36 km^2 with various levels of urbanization where urban changes are visible. Given the target application, only pairs of bitemporal images are provided.

Differently from standard land cover classification tasks, where the class types can be identified by analyzing single-time images (e.g., build-up, grassland, rivers), the discrimination of similar crop types is not trivial and requires the exploitation of temporal information. To create labeled data the know-how of expert annotators that analyze by photo-interpretation the whole TS of images is required. The alternative collection of *in situ* measurement is unfeasible at large scale. Hence, the scarcity of large-scale labeled multitemporal datasets for crop type mapping is due to the complexity of the considered task. In [7], a benchmark dataset for the supervised classification of field crops is proposed. The dataset, which is made up of 610 000 labeled instances, is located in the Brittany region, France. To generate the dataset, the publicly available crop type map of the considered region has been used. The map, based on the farmer's declarations, is released by the National Institute of Forest and Geography Information (IGN) in the context of Agricultural Land Parcel Information System (LIPS)—Registre Parcellaire Graphique (RPG). For each crop the average value of all the pixels associated to a given field parcel is computed per spectral band. The dataset, composed of Sentinel 2 images acquired between January 1, 2017 and December 31, 2017, provides the data in two processing levels: the raw reflectances at the top-of-atmosphere (level 1 C) and the atmospherically corrected surface reflectances at the bottom-of-atmosphere (level 2 A). The classification scheme is made up of nine crop types: "barley," "wheat," "rapeseed," "corn," "sunflower," "orchards," "nuts," "permanent meadows," and "temporary meadows." Although this crop-type benchmark is the very first that allows testing deep learning models for crop type mapping, the training set is made up of $\sim$ 319.000 labeled units, which are still a relatively small number for training deep learning architectures made up of several hidden layers and a large number of parameters. Moreover, the considered region is quite small compared to a standard country scale, i.e., 27.208 km^2. Finally, the extraction of the crop type labels directly from the map may lead to noisy labeled units, which are not correctly associated to their labels.

In this article, with the aim of advancing the capacity of exploiting Sentinel 2 TS for agricultural monitoring, we present a large-scale dataset called TimeSen2Crop that is made up of about 1 million of labeled samples belonging to 16 crop types. The dataset includes atmospherically corrected TSs of Sentinel 2 images collected in the agronomic year (i.e., period from one year's harvest to the next one for agricultural commodity) ranging between September 2017 and August 2018. The labeled units were collected in the full Austria, which extends for 83.879 km^2. The shadow, cloud, and snow pixels masks are provided for each labeled pixel. To generate the dataset, we considered the publicly available Austrian crop type map, which is based on farmer's declarations collected by surveys within the subsidy application process in the context of the common agricultural policy (CAP). Although the reliability of such product is very high, the direct extraction of labeled units from the map may lead to the selection of pixels:

1) located on the crop boundary (i.e., mixed spectral signature);
2) associated to crop type labels that are not correct for the whole year (rotation practice);
3) associated to the wrong crop type due to the spatial aggregation (polygon label represents the dominant class but is not correctly associated to all the pixels).

To solve this problem, the map-labeled units are selected according to an automatic training set extraction procedure based on a deep learning long short-term memory (LSTM) model. This multitemporal deep learning architecture is able to properly capture the phenological trend of the different cultivation. Therefore, it is reasonable to assume that the labeled units classified with the highest confidence are the pure pixels having high probability to be correctly associated to their labels.

To assess the quality of the resulting dataset and highlight its classification difficulties, a benchmarking experimental analysis has been carried out by comparing deep learning models employed in the literature for crop type mapping. Moreover, this article proposes a domain adaptation challenge, by providing the TS of Sentinel 2 images acquired in the agronomic year right after the one considered in the dataset (i.e., between September 2018 and August 2019). The goal of the challenge, which is related to the solution of a domain adaptation problem, is to direct more attention to one of the main critical issue of the crop classification task, which is the need of regularly updating the crop type maps, which strongly vary in the years due to the crop rotation practices.

The main contributions of this article can be summarized as follows. First, a novel labeled large-scale (about 1 million of samples) dataset based on multitemporal Sentinel 2 images for crop type mapping is proposed. The dataset has been generated by using an automatic system architecture, which extracts annotated samples having high reliability from the considered thematic product. Second, the practicality and the quality of the dataset are verified by testing and comparing several deep learning models employed for crop type mapping. Third, a challenge covering a typical domain adaptation problem related to crop type mapping is identified and the related dataset is released to allow the RS community to effectively address it in future studies.

The remainder of this article is organized as follows. The procedures employed to design the benchmark dataset are described in Section II. Section III gives an overview on the peculiar properties of the generated benchmark dataset, while Section IV introduces some deep learning models, used in the literature for crop type mapping, which are tested and compared on the proposed dataset. Section V illustrates the three proposed challenges related to the presented dataset. Finally, Section VI concludes this article.

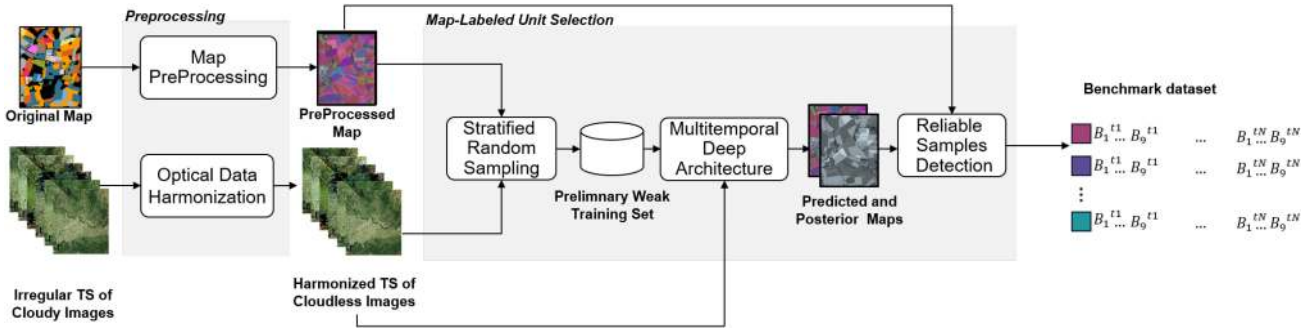

Fig. 1. Architecture of the automatic method for extracting reliable crop type labels from the public available agricultural Austrian thematic product.

II. Design of the TimeSen2Crop

In this section, we present the procedures employed to generate the TimeSen2Crop dataset by describing in detail the properties of the considered RS data and the thematic products used to generate the benchmark.

A. Sentinel 2 Data Collection

The considered agricultural study area is located in Austria, which is characterized by a complex landscape typical of the Alpine region. The topography of this country, which ranges from high mountain to lowlands areas, together with its peculiar climate (climatic gradient from west to east) lead to high biodiversity [8]. This results in a large variety of different agricultural landscape types, which occupies one fifth of the Austrian territory (i.e., 83.879 km^2).

Sentinel 2 data acquired between September 1, 2017 and September 1, 2018 were collected by discarding only the data having cloud coverage lower than 80%. The data were downloaded from the Food Security Thematic Exploitation Platform [9], where Sentinel 2 images made up of nine spectral bands are provided at 10 m spatial resolution, each having a size of 10980×10980 pixels. In particular, the blue (B2–490 nm), green (B3–560 nm), red (B4–665 nm), the four vegetation red edge (B5–705 nm, B6–740 nm, B7–0.783 nm, and B8A–865 nm) and the two short wave infrared (SWIR) (B11–1610 nm and B12–2190 nm) channels were considered. Band 8 was discarded because of its coarser spectral resolution compared to band 8 A. The data are atmospherically corrected using the radiative transfer model MODTRAN [10] and the spectral bands are provided with a spatial resolution of 10 m [11].

B. Sample Labeling

To provide a reliable benchmark dataset, an automatic machine-learning-based procedure has been considered to extract labeled units with high reliability from the considered publicly available Austrian crop type map [12]. This thematic product has been produced in the context of the CAP of the European Union, in order to verify eligibility for area-based subsidies. Hence, the crop types are based on farmer declarations, while the polygon field boundaries are the ones provided by the Land Parcel Identification System (LIPS) [13]. Although the reliability of such product is very high, the random selection of labeled units extracted directly from the map may lead to the following:

1) outdated labeled samples;
2) crop types labels not valid for the whole agronomic year due to the rotation practice;
3) samples associated to the wrong labels due to the polygon spatial aggregation.

Hence, even though the map represents a rich information source, it is necessary to accurately select the labeled units to detect spectral pixels correctly associated to their labels [14]. Fig. 1 shows the automatic approach employed to address this issue, which is made up on the following two main phases: 1) preprocessing, and 2) map-labeled unit selection.

1) Preprocessing: In this step, we apply a preprocessing to both the optical data and the thematic product. The optical preprocessing step aims to the following: 1) spatially and temporally harmonize the irregular TSs of cloudy images, and 2) mitigate the cloud coverage problems. For each pixel associated to a labeled unit, we extract the multitemporal spectral feature vector. Such vector, which represents the whole TS of images, is converted into a TS of 12 monthly composites using a pixel-based statistic-based approach.

Let us focus the attention on the set of Q Sentinel 2 images acquired within the ith month, with $i \in [1, 2, \ldots, 12]$. Let $\mathbf{X}_j^i = [\mathbf{x}_j^{i,1}; \mathbf{x}_j^{i,2}; \cdots; \mathbf{x}_j^{i,Q}]^T$ be the multitemporal spectral matrix associated to the jth labeled pixel of the dataset, where $\mathbf{x}_j^{i,1} = [x_{j,1}^{i,1}, x_{j,2}^{i,1}, \ldots, x_{j,N}^{i,1}]$ is the spectral vector of the first Sentinel 2 image of the considered ith month made up of $N = 9$ spectral channels, i.e., $\mathbf{X}_j^i \in \mathbb{R}^{N \times Q}$. For each spectral band, the Q reflectance values are collapsed into a single one (representing the month) by computing their median. Let $\mathbb{M}\{\bullet\}$ be the median operator, the computation of the ith monthly composite is as follows:

$$\begin{aligned} \hat{x}_{j,1}^i &= \mathbb{M}\{x_{j,1}^{i,1}, x_{j,1}^{i,2}, \ldots, x_{j,1}^{i,Q}\} \\ \hat{x}_{j,2}^i &= \mathbb{M}\{x_{j,2}^{i,1}, x_{j,2}^{i,2}, \ldots, x_{j,2}^{i,Q}\} \\ &\vdots \\ \hat{x}_{j,N}^i &= \mathbb{M}\{x_{j,B}^{i,1}, x_{j,B}^{i,2}, \ldots, x_{j,N}^{i,Q}\} \end{aligned} \tag{1}$$

Fig. 2. Qualitative examples of map labeled units associated to the wrong crop category: (a) original crop type map, (b) high resolution aerial image, and (c) preliminary predicted map. The administrative boundary of the crop field do not match perfectly the real cultivation present in the scene. Applying the proposed automatic procedure, this problem is highly mitigated.

where $\hat{\mathbf{x}}_j^i = [\hat{x}_{j,1}^i, \hat{x}_{j,2}^i, \ldots, \hat{x}_{j,N}^i]$ is the spectral vector of the ith monthly composite. At the end of this step, the obtained multitemporal spectral vector $[\hat{\mathbf{x}}_j^1, \hat{\mathbf{x}}_j^2, \ldots, \hat{\mathbf{x}}_j^{12}]$ is made up of 108 features, i.e., 9 reflectance values $\times$ 12 months. The median computation is performed by discarding the cloudy, snowy, and shadowy samples. If no cloud-free images are available for a considered pixel within the month, the harmonization process set all the reflectance values of the month to zero.

The thematic product is preprocessed in order to 1) convert the map legend into the desired classification scheme, and 2) remove the pixels close to the boundary of the crop fields, which may be related to spectrally mixed pixels due to the spatial resolution of Sentinel 2. The map legend has been carefully revised by RS and agricultural experts to define a set of crop categories interesting from the agricultural view point and discriminable according to the spectral and temporal information provided by Sentinel 2. Details are provided in Section III-A. To remove the pixel close to the field boundary, the standard erosion morphological operator having disk element of radius equal to 3 is applied to the thematic map [15]. This step highly increases the probability of selecting samples having almost pure spectral signature, i.e., representative of the phenological properties of a single crop type.

2) Selection of Map-Labeled Units: To generate a crop type mapping dataset, the prior probabilities of the crop categories are extracted from the thematic product using a stratified random sampling approach. This sampling strategy extracts a number of samples per crop type proportional to the number of crops associated with that type in the considered study area. This procedure allows us to generate a preliminary "weak" training set where misclassified samples may be present (i.e., samples associated to wrong labels) [14]. Then, we run the classifier to generate a preliminary predicted map by using a multitemporal deep learning LSTM network, a special kind of RNN widely used for crop type mapping [16]. This network has been extensively employed to elicit temporal patterns due to its long-term memory capabilities. By storing a huge amount of evidence to make decisions in that actual temporal context, it provides better solutions compared to other recurrent deep learning models. The predicted map is compared to the original one for selecting only the samples located in the areas where the two maps agree. Moreover, the pixel wise posterior probabilities provided by the network are used as a confidence measure of the classification result. Only the pixels having high probability to be correctly classified are selected. In particular, the samples having pixelwise posterior probabilities higher than the 75th percentile of all posteriors of that class are selected as possible candidates. This rule allows us to select the samples classified with high-confidence, adaptively computing a different threshold value per class. To capture the spatial variability of the crop classes all over the study area and reducing correlation, we impose the constrain that the Sentinel 2 pixel labeled units included in the benchmark dataset must have a distance of at least 12 pixels (i.e., 120 m) each others.

Fig. 2 shows a qualitative example of map labeled units associated to the wrong crop category. The administrative boundaries of the crop field do no match perfectly the real cultivations present in the scene as visible in Fig. 2(a) and (b), where the original crop type map and the high-resolution aerial image are reported. Applying the proposed procedure, this problem is mitigated [see Fig. 2(c)], thus highly increasing the probability of associating the map label to pure spectral samples in the TSs of Sentinel 2 images.

III. TimeSen2Crop Properties

The Austrian country is covered by the 15 Sentinel 2 tiles shown in Fig. 3. Samples extracted from spatially disjoint tiles were included into the training set (13 tiles), the test set (1 tile), and the validation set (1 tile) to generate three statistically independent sets, each with 822 843 (76.71%), 133 419 (12.43%), and 116 369 (10.84%) labeled pixels, respectively. Please note that this condition allows us to have a test area, which extends for 12056.4 km^2 and is statistically independent from the training ones since no spatial overlapping is considered between the training and the test samples. Fig. 3 shows the division into training, test, and validation of the Sentinel 2 tiles.

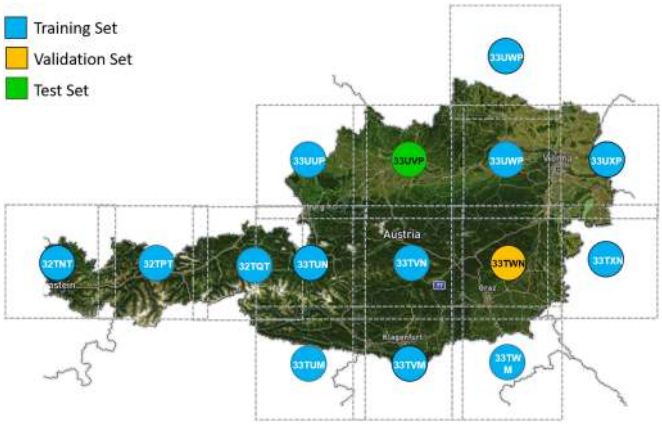

Fig. 3. Spatial division of the Sentinel 2 tiles into training, test and validation is reported in blue, yellow, and green, respectively.

TABLE I
CLASS DISTRIBUTIONS IN THE TIMESEN2CROP DATASET

Class Name	N. of Samples	Prior Prob.(%)
Legumes	7098	0.66
Grassland	263202	24.53
Maize	144315	13.45
Potato	27768	2.28
Sunflower	22636	2.11
Soy	61134	5.69
Winter Barley	74060	6.90
Winter Caraway	1270	0.11
Rye	46191	4.30
Rapeseed	37978	3.54
Beet	30491	2.84
Spring Cereals	77242	7.20
Winter Wheat	112326	10.47
Winter Triticale	54106	5.04
Permanent Plantation	46102	4.29
Other Crops	66712	6.21

A. Classification Scheme

The definition of the classification scheme is a fundamental step for the proper generation of a reliable and consistent benchmark dataset. In particular, the classification scheme must include only classes that can be discriminated by using the multitemporal multispectral information provided by the TSs of Sentinel-2 images. To this end, the map legend of the considered thematic product [12] was carefully revised by RS and agricultural expert to fill the semantic gap between the crop categories present in the map and the RS data. In greater detail, the proposed classification scheme includes the 16 crop categories reported in Table I. Fig. 4 presented the distribution of the crop categories per tile, while Table I reports the pixel count (i.e., number of labeled units) and the proportion per crop category.

The "grassland," "spring cereals," "legumes," and "permanent plantations" crop categories semantically aggregate minor classes, which cannot be distinguished from the phenological view point using multispectral optical images. The "grassland" is made up of "clover" (52.59%), "green pruning rye" (0.15%), and "alpine meadow" (47.24%), while the "spring cereals" includes "spring oat" (30.94%), "spring wheat" (14.52%), and "spring barley" (54.52%). The "legumes" includes horsebeans (97.80%) and a small presence of sweet lupines (2.19%). The "permanent plantations" crop type includes all the following fruit trees: Vineyards (83.29%), Cherry Plantation (0.47%), Apricots (1.48%), Nectarines (0.02%), Peach (0.44%), Apples (0.94%), Pears (12.89%), and Plums (0.44%). Finally, "other crops" includes all the remaining minor classes present in the map to have an exhaustive classifications scheme. Due to their scarcity, these crops are not enough to represent a singular crop type. However, the presence of such a class is fundamental to consider that not all the crops can be modeled and represented in the training set but an exhaustive classification scheme is required in real-world crop type mapping problems.

Differently from the RS dataset publicly available, the proposed benchmark dataset is characterized by a detailed classification scheme that leads to a challenging crop type mapping problem.

B. Temporal Properties

The proposed benchmark is made up of TSs of images acquired in the agronomic year ranging between September 2017 and August 2018. Differently from the standard yearly-based TSs, the provided temporal information allows the accurate characterization of the phenological trend of the cultivations [17]. Only Sentinel 2 images having cloud coverage $< 80\%$ are included in the dataset. Although such threshold is quite conservative, many images acquired in winter have been discarded from the TSs since the considered study area is affected by heavy cloud coverage.

Table II reports the acquisition dates of the images included in the TSs for each Sentinel 2 tile. One can notice that TSs acquired over different tiles present 1) unequal lengths, and 2) variations in the temporal sampling rate. Indeed, due to the heavy cloud and snow coverage, for some tiles no images are available for some months, thus affecting their temporal sampling. Moreover, even though Sentinel 2 satellites are characterized by a large swath (i.e., 290 km width), different along-track strips are required to cover large study areas. This leads to TSs:

1) having variations between the acquisition time of the first and the last image of the TSs;
2) made up of images acquired in different time stamps (different temporal sampling);
3) having sequences with variable lengths.

These issues are well known at operational level when working at country or continental scale. Therefore, the temporal properties of the proposed challenging benchmark dataset accurately depict real-world scenarios.

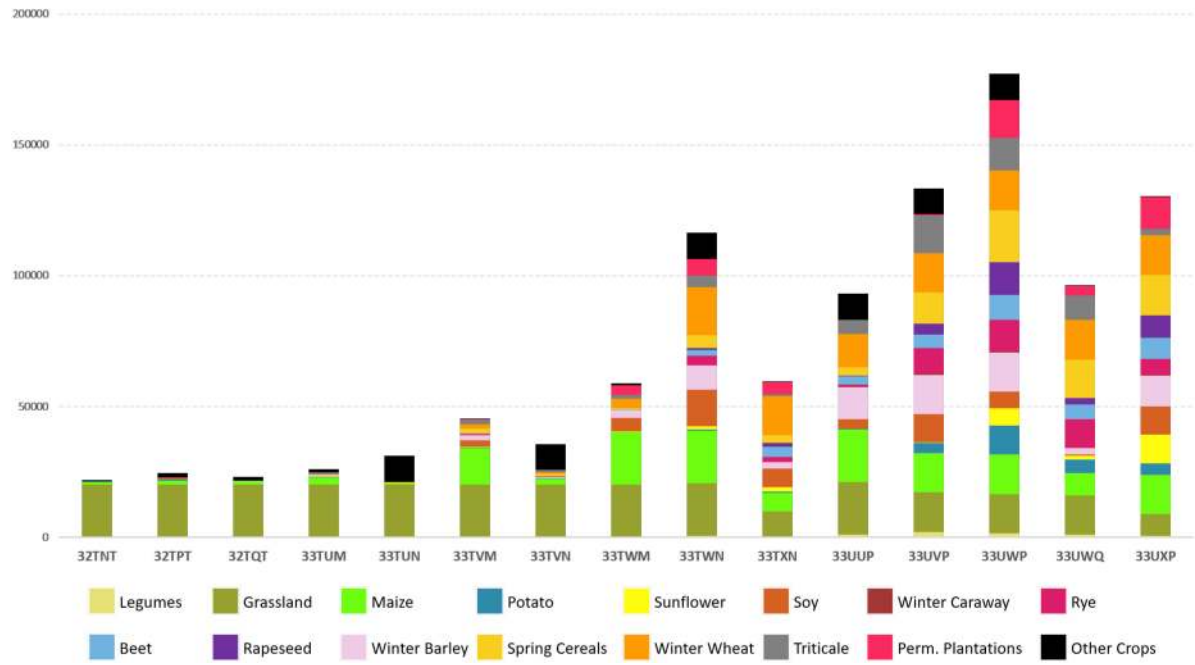

Fig. 4. Distributions of the crop categories of the TimeSen2Crop dataset per tile.

C. Structure of the Dataset

TimeSen2Crop aims to contribute to the worldwide research related to the classification of TSs of Sentinel 2 images for crop type mapping. Thus, it is publicly available at the website TimeSen2Crop[1]. The structure of the dataset is hierarchical. The data are first organized per Sentinel 2 tile, in 15 folders. Each Sentinel 2 folder includes 16 subfolders, i.e., the crop categories and the acquisition dates of the specific tile (see Table II) reported in a CSV file. The dates are time-ordered from the oldest acquisition to the newest.

Inside the crop type subfolders, each labeled sample has associated a CSV file, which stores the multitemporal spectral signature of the corresponding Sentinel 2 pixel. The number of files depends on the number of pixels extracted for the considered crop category for the specific Sentinel 2 tile. The spectral values of the pixels are stored in a matrix, which indicates the reflectance values for each spectral channel within the different acquisition dates. In particular, each row reports the different acquisition time, while each column constrains the different spectral bands, namely, the blue (B2), green (B3), red (B4), the four vegetation red edge (B5, B6, B7, and B8A), and the two SWIR (B11 and B12). The last column contains the information regarding the condition of the pixel, which can be clear (value 0), cloud (value 1), shadow (value 2), or snow (value 3). Note that as the considered study area is heavily affected by snow coverage during the winter season, it is important to have precise information on the snow presence in the scene.

D. Data Imbalance

Real-world crops classification tasks are usually characterized by classes having severely imbalanced prior probabilities. Agricultural areas are usually dominated by few common crops, which correspond to the crops cultivated extensively (e.g., corn, wheat). This leads to very different prior probabilities of the various classes present in the scene. When dealing with such high imbalanced datasets, the classification model may fail to accurately recognize the minority classes. However, the accurate classification of minority crop types is still important. Due to the use of a stratified random sampling strategy and the possibility of having available the thematic product for the whole Austrian country, the proposed benchmark dataset is a valid example of challenging crop type mapping datasets since it is affected by a real and strong class imbalance. Thus, it can be used to assess the capability of different techniques to address this common problem in agriculture applications.

[1][Online]. Available: https://rslab.disi.unitn.it/timesen2crop/

IV. Benchmarking Algorithm Comparison

To assess the complexity of the proposed TimeSen2Crop dataset, we have evaluated and compared existing deep learning methods typically employed for crop type mapping training them from scratch. We focus the attention on deep learning models since they outperformed shallows models such as random forest or support vector machine [18]. Moreover, the proposed large-scale dataset is meant to be used for deep architectures that require a large number of samples to successfully train the model [19]. This peculiar classification task has been addressed considering two main categories of multitemporal deep learning architectures, i.e., recurrent deep learning models and time-convolution deep learning models. While recurrent models are suited for crop type mapping due to their capability of capturing temporal dynamics, time-convolution models directly focus on the temporal profiles to accurately classify sequential data.

TABLE II
IMAGE ACQUISITION DATES (I.E., MONTH/DAY) OF THE TSS ACQUIRED OVER THE DIFFERENT TILES ARE REPORTED

	32TNT	32TPT	32TQT	33TUM	33TUN	33TVM	33TVN	33TWM	33TWN	33TXN	33UUP	33UVP	33UWP	33UWQ	33UXP
2017	9/4	9/4	9/21	9/21	9/26	9/3	9/8	9/13	9/13	9/5	10/11	9/8	9/8	9/18	9/5
	9/24	10/14	9/26	10/6	10/11	9/13	9/23	9/18	9/23	9/10	10/16	9/18	9/18	9/28	9/15
	10/4	10/19	10/6	10/11	10/16	9/23	9/28	9/23	9/28	9/30	10/26	9/23	9/23	10/13	9/30
	10/14	10/24	10/11	10/16	10/26	9/28	10/18	9/28	10/13	10/5	10/31	9/28	9/28	11/2	10/15
	10/19	11/3	10/16	10/26	10/31	10/8	12/2	10/13	10/18	10/15	11/5	11/2	10/13	11/22	10/30
	10/24	11/18	10/26	10/31	11/15	10/13	-	10/18	11/27	10/30	-	12/7	10/18	12/7	11/4
	11/3	12/3	10/31	11/10	12/25	10/18	-	11/27	12/2	11/4	-	-	11/2	-	11/14
	12/3	12/13	11/15	11/15	-	11/27	-	12/2	12/17	11/14	-	-	11/27	-	12/19
	12/13	-	12/25	12/5	-	12/7	-	12/7	-	12/19	-	-	12/2	-	12/24
	-	-	-	12/20	-	12/17	-	12/17	-	-	-	-	12/7	-	12/29
	-	-	-	12/25	-	12/22	-	-	-	-		-	-	-	-
2018	1/27	1/12	1/14	1/24	1/24	1/21	1/6	1/6	2/25	3/24	1/19	1/6	1/6	1/6	3/24
	2/26	1/27	1/24	1/29	1/29	2/15	1/31	1/21	3/22	4/8	1/24	2/5	2/15	2/15	4/8
	3/8	2/26	1/29	2/28	2/13	2/25	2/5	2/15	4/6	4/13	1/29	2/25	2/25	2/25	4/13
	4/2	3/8	2/13	3/20	2/28	3/22	2/25	3/7	4/21	4/18	2/13	3/2	3/7	3/22	4/18
	4/7	3/13	2/23	3/25	3/5	4/1	3/22	3/12	5/6	4/23	2/28	3/22	3/22	4/6	4/28
	4/17	4/2	2/28	4/14	3/25	4/6	4/6	3/22	5/31	4/28	3/15	4/6	4/6	4/11	5/8
	4/22	4/7	3/5	4/19	3/30	4/11	4/11	4/6	6/10	5/8	3/25	4/11	4/11	4/21	5/13
	4/27	4/17	3/25	4/24	4/4	4/21	4/21	4/11	6/20	5/13	3/30	4/21	4/21	5/1	6/12
	5/7	4/22	3/30	4/29	4/9	5/6	5/6	4/21	7/20	6/2	4/4	5/6	5/1	5/6	6/17
	5/12	4/27	4/14	5/9	4/14	5/26	5/26	5/6	7/25	6/12	4/14	5/21	5/6	5/21	6/27
	6/1	5/7	4/19	5/19	4/19	5/31	5/31	5/21	7/30	6/17	4/19	5/26	5/21	5/26	7/2
	6/6	6/6	4/24	6/3	4/24	6/5	6/5	5/26	8/4	6/27	4/24	5/31	5/26	5/31	7/7
	6/16	6/16	4/29	6/8	4/29	6/10	6/15	5/31	8/9	7/2	4/29	6/5	5/31	6/5	7/12
	6/21	6/21	5/14	6/18	5/14	6/15	6/20	6/5	8/19	7/7	5/9	6/10	6/5	6/10	7/17
	6/26	7/1	5/19	6/23	5/19	6/20	7/15	6/10	8/29	7/17	5/14	6/20	6/10	6/15	7/27
	7/1	7/11	6/3	7/3	6/3	6/30	7/20	6/20	-	7/27	5/19	6/30	6/20	6/20	8/1
	7/11	7/16	6/8	7/13	6/8	7/5	7/25	6/30	-	8/1	5/29	7/5	6/30	6/30	8/6
	7/16	7/26	6/23	7/18	6/23	7/10	7/30	7/5	-	8/11	6/3	7/15	7/5	7/5	8/16
	7/26	7/31	7/3	7/23	7/3	7/20	8/4	7/10	-	8/16	6/8	7/20	7/15	7/15	8/21
	7/31	8/5	7/13	7/28	7/13	7/25	8/9	7/20	-	8/21	7/3	7/25	7/20	7/20	-
	8/5	8/15	7/18	8/2	7/18	7/30	8/19	7/25	-	8/31	7/8	7/30	7/25	7/25	-
	8/15	8/20	7/28	8/7	7/28	8/4	8/29	7/30	-	-	7/13	8/9	7/30	7/30	-
	8/20	-	8/2	8/12	8/2	8/9	-	8/4	-	-	7/18	8/19	8/4	8/4	-
	-	-	8/7	8/17	8/7	8/19	-	8/9	-	-	7/23	8/29	8/9	8/9	-
	-	-	8/12	8/22	8/12	8/24	-	8/19	-	-	7/28	-	8/19	8/19	-
	-	-	8/17	8/27	8/17	-	-	8/24	-	-	8/2	-	8/29	8/29	-
	-	-	8/22	-	8/22	-	-	8/29	-	-	8/7	-	-	-	-
	-	-	8/27	-	8/27	-	-	-	-	-	8/12	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	8/17	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	8/22	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	8/27	-	-	-	-

Only Sentinel 2 images having cloud coverage $< 80\%$ are included in the dataset.

The benchmark time-convolution deep learning models considered are InceptionTime [20], Multiscale ResNet (MSResNet) [21] and temporal convolutional neural network (TempCNN) [22], while the recurrent deep learning models selected are the LSTM network [16], the STAR recurrent neural network (StarRNN) [24], and a weighted LSTM suited for imbalanced data problems. In particular, the weighted LSTM has been trained by the two-step procedure proposed in [25], which mitigates the problem of having imbalanced training data. In the first step, the weights of the LSTM cost function per crop type are set according to the number of samples of each class under the assumption that the number of samples are in proportion to the a priori probability of the different crops. This procedure avoids that the gradient calculated with respect to the network weights is dominated by the contributions of the dominant classes. Then, the network weights obtained at the end of the first phase are used as the initial ones of the second training phase, which is performed using the standard cost function. This operation allows the output to be restored as an approximation of the *a posteriori* probabilities. Finally, a self-attention transformer model, originally developed as sequence-to-sequence encoder–decoder models for language translation [26], was tested on the proposed dataset. This attention-based method has been selected since it proved to be effective for crop type mapping problems [18], [23].

The experimental analysis has been carried out considering the trainSet, testSet, and validationSet of the proposed dataset. Since all the considered deep learning models assume to deal with homogeneous TSs characterized by uniform length, the optical preprocessing step presented in Section II-B1 is first applied to harmonize the TSs acquired over different tiles. To compare the performance obtained by the different deep learning models, the Fscore (F1%) and the overall accuracy (OA%) metrics are evaluated on the test set. The standard grid search approach was used to train the models. The

TABLE III
OA% AND THE F1% OBTAINED ON THE TEST SET

Classes	Time-based Convolution Models			Attention-based Models	Recurrence-based Models		
	Inc. Time	MSResNet	TempCNN	Transformer	StarRNN	LSTM	LSTM Weig.
Legumes	76.67	80.93	63.69	80.25	71.85	80.26	**83.23**
Grassland	79.50	85.24	82.04	**85.82**	84.75	82.90	82.99
Maize	97.57	97.89	97.87	97.50	95.76	97.56	**97.94**
Potato	83.75	**89.33**	74.62	80.33	83.06	82.99	84.93
Sunflower	67.16	**75.12**	61.33	60.85	40.05	59.96	70.89
Soy	91.42	**97.27**	93.54	94.57	93.40	95.29	96.15
Winter Barley	**93.27**	90.48	85.99	92.27	84.39	88.42	91.92
Winter Caraway	33.52	**76.24**	0.00	65.14	24.76	50.81	42.95
Rye	76.11	72.37	68.36	**76.66**	67.03	69.76	73.96
Rapeseed	97.53	96.29	94.09	**98.41**	97.36	97.29	96.98
Beet	95.25	94.35	96.61	93.94	93.08	92.66	**97.04**
Spring Barley	84.63	87.44	85.35	87.10	86.47	**89.62**	89.49
Winter Wheat	92.94	75.41	94.15	93.24	91.91	94.08	**95.44**
Winter Triticale	67.34	67.71	67.38	71.19	62.17	69.79	**75.80**
Perm. Plantation	**39.64**	25.42	23.92	29.61	26.13	34.00	26.52
Other Crops	44.57	**48.45**	22.15	42.05	26.06	36.23	45.23
OA%	81.90	81.80	81.71	84.44	81.17	83.44	**85.39**
Median F1%	81.62	83.09	78.33	83.07	83.73	82.95	**84.08**

The scores are reported for the following benchmark methods: InceptionTime [20], MSResNet [21], TempCNN [22], attention-based Transformer [23], StarRNN [24], LSTM [16], and weighted LSTM.

optimal setup is identified according to the accuracy achieved on the validationSet. The model specific parameters tested varied per method according to what suggested by the authors. The learning rate and the weight decay were sampled from log-uniform distributions, respectively $U_{\log}([10^{-2}, 10^{-4}])$ and $U_{\log}([10^{-2}, 10^{-8}])$. The weighted LSTM, the LSTM and the StarRNN have been analyzed with different cascaded layers $L \in \{2, 3, 4\}$ and hidden representation $H \in \{2^5, 2^6, 2^7\}$. The Inception model has been evaluated with different stacks of Inception modules $L \in \{1, 2, 3, 4\}$ and hidden representations $H \in \{2^5, 2^6, 2^7\}$. The MSResNet model hidden representations were selected within $H \in \{2^5, 2^6, 2^7, 2^8, 2^9\}$. The kernel size and the hidden representations for the TempCNN were searched over $K \in \{3, 5, 7\}$ and $H \in \{2^5, 2^6, 2^7\}$, respectively. Lastly, the Attention model was validated with $N_{\text{head}} \in \{1, 2, \ldots, 8\}$ by using $L \in \{1, 2, \ldots, 8\}$ stacked layers and $H \in \{2^5, 2^6, 2^7\}$ hidden representations. Please note that no extensive tuning of the hyperparameters has been performed, since the goal of the proposed benchmark comparison is to assess the quality of the proposed dataset.

Table III reports the experimental results obtained per method. As expected the most critical classes are the "permanent plantations" and "other crops" regardless of the deep learning model. This is due to the semantic aggregation of these classes, which include several types of cultivations. The results obtained show that the considered benchmark methods achieve similar performances, with the OA% ranging from a minimum of 81.71% (TempCNN) to a maximum of 85.39% (weighted LSTM). In general, the time-convolution models performed a little worst compared to the LSTM and the attention models. Hence, InceptionTime, MSResNet, and TempCNN obtain an OA of 81.90%, 81.80%, 81.71%, respectively, compared to of 85.39%, 83.44%, and 84.44% of the weighted LSTM, LSTM, and the Transformer models, respectively. As expected the weighted LSTM obtains better results with respect to the standard LSTM due to its capability of better handling the imbalanced problem. Moreover, this method is the one that achieve the highest median F1% equal to 84.08%. These results point out the main properties of the TimeSen2Crop dataset and highlights the related classification difficulties.

V. PROPOSED TIMESEN2CROP CHALLENGE

In this section, we focus on one of the main challenge of real-world crop type mapping: exploit trained deep learning models to classify a TS of Sentinel 2 images acquired in a different agronomic year. This is a particular problem of domain adaptation. In particular, we focus the attention on the TSs acquired in an agronomic year after the one considered in the dataset, i.e., on images acquired between September 2018 and August 2019. For making it possible the evaluation of the performance of a given method under this critical condition, a testSet is distributed that has been acquired over the 33UVP Sentinel 2 tile. Similar to the original testSet of the TimeSen2Crop dataset, the images included in the TS present a cloud coverage smaller than 80%. Also in this case, information about shadow, clouds, and snow are reported for each labeled samples. This challenge opens to the possibility to test the capability of adapting different multitemporal deep learning models to different year acquisition (domain adaptation) assuming that no reference data are available for the new year. Note that one of the most important problems from the operational view point is the production of consistent land-cover maps for multiple years. To overcome this problem, sequential cascade classification methods [27],

[28] and compounds methods [29] have been proposed in the literature. They typically rely the temporal correlation between TSs acquired over the same area in the different years. Moreover, large effort has been devoted in the literature for adapting pretrained deep learning models to different datasets sharing similar properties considering fine-tuning strategies [30], [31]. However, also in this case, crop type mapping represents a peculiar classification task since most of the crops change from one year to the other because of the crop rotation practice. Moreover, no comprehensive analyses have been carried out to solve the problem of generating multiple years crop type maps.

VI. Conclusion

This article presents a multitemporal benchmark dataset based on TSs of atmospherically corrected Sentinel 2 images for large-scale crop type mapping called TimeSen2Crop. The dataset presents a detailed classification scheme made up of 16 crop categories including $\sim$1 million of labeled units collected in the Austrian country, which extends for 83.879 km^2. For each labeled sample, the cloud, shadow, and snow information are provided. The benchmark is provided divided into three spatially disjoint sets for training the models (trainSet), optimizing the parameters (validationSet), and evaluating the performances (testSet). Several benchmarking algorithms and experimental results have been evaluated and compared on the proposed dataset. The experimental results point out the properties of the dataset and provide an overview of its classification difficulties. Moreover, TimeSen2Crop proposes a domain adaptation challenge to address one of the main critical issues of crop type mapping, which is the need of frequently update the maps due to the crop rotation practice. TimeSen2Crop is publicly available to allow the community to develop advanced methods for crop type mapping at large scale. The dataset can be found at TimeSen2Crop[2].

Acknowledgment

The authors would like to thank the VISTA Remote Sensing in Geosciences GmbH partner for providing the atmospherically corrected Sentinel 2 images in the Food Security TEP platform. This benchmark dataset has been developed in the framework of the ExtremeEarth Project.

References

[1] 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, Jun. 2019.

[2] I. Demir *et al.*, "Deepglobe 2018: A challenge to parse the earth through satellite images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 172–209.

[3] 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," 2019.

[4] 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.*, Jul. 2019, pp. 5901–5904.

[5] B. Bischke *et al.*, "The multimedia satellite task at mediaeval 2019: Estimation of flood severity," in *Proc. CEUR Workshop*, 2019.

[6] R. C. Daudt, B. Le Saux, A. Boulch, and Y. Gousseau, "Urban change detection for multispectral earth observation using convolutional neural networks," in *Proc. IEEE Int. Geosci Remote Sens. Symp.*, 2018, pp. 2115–2118.

[7] M. Rußwurm, C. Pelletier, M. Zollner, S. Lefèvre, and M. Körner, "BreizhCrops: A time series dataset for crop type mapping," 2019, *arXiv:1905.11893*.

[8] I. Schmitzberger, T. Wrbka, B. Steurer, G. Aschenbrenner, J. Peterseil, and H. Zechmeister, "How farming styles influence biodiversity maintenance in austrian agricultural landscapes," *Agriculture*, *Ecosystems Environ.*, vol. 108, no. 3, pp. 274–290, Jun. 2005.

[9] M. Muerth *et al.*, "Food security Tep-supporting sustainable intensification of food production from space," in *Proc. IOP Conf. Ser.: Earth Environ. Sci.*, vol. 509, no. 1, Jun. 2020, Art. no. 0 12038.

[10] A. Berk *et al.* "MODTRAN4 radiative transfer modeling for atmospheric correction," in *Proc.* SPIE, vol. 3756, pp. 348–353, 1999.

[11] W. Verhoef and H. Bach, "Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models," *Remote Sens. Environ.*, vol. 87, no. 1, pp. 23–41, 2003.

[12] 2019, Accessed: Sep. 10, 2020: [Online]. Available: https://www.data.gv.at/katalog/dataset/e21a731f-9e08-4dd3-b9e5-cd460438a5d9

[13] K. Kocur-Bera, "Data compatibility between the land and building cadaster (LBC) and the land parcel identification system (LPIS) in the context of area-based payments: A case study in the polish region of warmia and mazury," *Land Use Policy*, vol. 80, pp. 370–379, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0264837718300711

[14] C. Paris and L. Bruzzone, "A novel approach to the unsupervised extraction of reliable training samples from thematic products," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 1930–1948, Mar. 2021.

[15] P. Soille, *Morphological Image Analysis: Principles and Applications*. Berlin, Germany: Springer, Mar. 2013.

[16] M. Rußwurm and M. Korner, "Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2017, pp. 11–19.

[17] Y. T. Solano-Correa, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Spatio-temporal evolution of crop fields in Sentinel-2 satellite image time series," in *Proc. 9th Int. Workshop Anal. Multitemporal Remote Sens. Images*, 2017, pp. 1–4.

[18] M. Rußwurm, S. Lefèvre, and M. Körner, "BreizhCrops: A satellite time series dataset for crop type identification," *CoRR*, vol 2019. [Online]. Available: http://arxiv.org/abs/1905.11893

[19] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state-of-the-art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[20] H. I. Fawaz *et al.*, "Inceptiontime: Finding AlexNet for time series classification," *Data Mining Knowl. Discovery*, vol. 34, pp. 1936–1962, 2020.

[21] F. Wang, J. Han, S. Zhang, X. He, and D. Huang, "CSI-Net: Unified body characterization and action recognition," 2018, *arXiv:1810.03064*.

[22] 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, p. 523, 2019.

[23] M. Rußwurm and M. Körner, "Self-attention for raw optical satellite time series classification," *ISPRS J. Photogrammetry Remote Sens.*, vol. 169, pp. 421–435, 2020.

[24] M. O. Turkoglu, S. D'Aronco, J. D. Wegner, and K. Schindler, "Gating revisited: Deep multi-layer RNNs that can be trained," 2019, *arXiv:1911.11033*.

[25] L. Bruzzone and S. B. Serpico, "Classification of imbalanced remote-sensing data by neural networks," *Pattern Recognit. Lett.*, vol. 18, no. 11–13, pp. 1323–1328, 1997.

[26] A. Vaswani *et al.*, "Attention is all you need," 2017, *arXiv:1706.03762*.

[27] L. Bruzzone and R. Cossu, "A multiple-cascade-classifier system for a robust and partially unsupervised updating of land-cover maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 1984–1996, Sep. 2002.

[28] L. Bruzzone and D. F. Prieto, "A partially unsupervised cascade classifier for the analysis of multitemporal remote-sensing images," *Pattern Recognit. Lett.*, vol. 23, no. 9, pp. 1063–1071, 2002.

[29] B. Demir, F. Bovolo, and L. Bruzzone, "Detection of land-cover transitions in multitemporal remote sensing images with active-learning-based compound classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1930–1941, May 2012.

[2][Online]. Available: https://rslab.disi.unitn.it/timesen2crop/

[30] J. Li, X. Huang, and J. Gong, "Deep neural network for remote-sensing image interpretation: Status and perspectives," *Nat. Sci. Rev.*, vol. 6, no. 6, pp. 1082–1086, 2019.
[31] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

Giulio Weikmann received the "Laurea" (B.S.) degree and the "Laurea Specialistica" (M.S.) (*summa cum laude*) degree in information and communication engineering from the University of Trento, Trento, Italy, where he is currently working toward the Ph.D. degree with the Remote Sensing Lab (RSLab).

He is currently a Teaching Assistant with the Department of Information Engineering and Computer Science, University of Trento. His major research interest concerns image and signal processing and the development of deep learning architectures for the automatic processing and classification of remote sensed optical data. He conducts research on these topics within the frameworks of national and international projects.

Claudia Paris (Member, IEEE) received the "Laurea" (B.S.) degree and the "Laurea Specialistica" (M.S.) (*summa cum laude*) degree in telecommunication engineering and the Ph.D. degree in information and communication technology from the University of Trento, Trento, Italy, in 2010, 2012, and 2016, respectively.

She accomplished the Honors Master Program in Research within the master's degree in telecommunication engineering in 2012. Since 2014, she is a Teaching Assistant with the Department of Information Engineering and Computer Science, University of Trento, where she is currently an Assistant Professor. Her main research includes image and signal processing, machine learning and deep learning with applications to remote sensing image analysis remote sensing single date and time series image classification, land cover map update, and fusion of multisource remote sensing data for the estimation of biophysical parameters. She conducts research on these topics within the frameworks of national and international projects.

Dr. Paris was the recipient of the very prestigious Symposium prize paper award (SPPA) at the 2016 International Symposium on Geoscience and Remote Sensing (Beijing, China, 2016) and at the 2017 International Symposium on Geoscience and Remote Sensing (Fort Worth, Texas, USA, 2017). She has been a member of the program and scientific committee of several international conferences and workshops. She is a Reviewer for many international journals, among them the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Lorenzo Bruzzone (Fellow, IEEE) received the Laurea (M.S.) degree in electronic engineering (*summa cum laude*) and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1993 and 1998, respectively.

He is currently a Full Professor of Telecommunications with the University of Trento, Trento, Italy, where he teaches remote sensing, radar, and digital communications. He is currently the Founder and the Director of the Remote Sensing Laboratory (https://rslab.disi.unitn.it/), Department of Information Engineering and Computer Science, University of Trento. His current research interests are in the areas of remote sensing, radar and SAR, signal processing, machine learning and pattern recognition. He promotes and supervises research on these topics within the frameworks of many national and international projects. He is the Principal Investigator of many research projects. Among the others, he is currently the Principal Investigator of the Radar for icy Moon exploration (RIME) instrument in the framework of the JUpiter ICy moons Explorer (JUICE) mission of the European Space Agency (ESA) and the Science Lead for the High Resolution Land Cover project in the framework of the Climate Change Initiative of ESA. He is the author (or coauthor) of 294 scientific publications in referred international journals (221 in IEEE journals), more than 340 papers in conference proceedings, and 22 book chapters. He is Editor/Coeditor of 18 books/conference proceedings and one scientific book. His papers are highly cited, as proven from the total number of citations (more than 37 000) and the value of the h-index (89) (source: Google Scholar).

Dr. Bruzzone was invited as keynote speaker in more than 40 international conferences and workshops. Since 2009, he has been a member of the Administrative Committee of the IEEE Geoscience and Remote Sensing Society (GRSS), where since 2019, he is a Vice-President for Professional Activities. He ranked first place in the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Seattle, July 1998. Since that he was the recipient of many international and national honors and awards, including the recent IEEE GRSS 2015 Outstanding Service Award, the 2017 and 2018 IEEE IGARSS Symposium Prize Paper Awards and the 2019 WHISPER Outstanding Paper Award. He was a Guest Co-Editor of many Special Issues of international journals. He is the co-founder of the IEEE International Workshop on the Analysis of Multi-Temporal Remote-Sensing Images (MultiTemp) series and is currently a member of the Permanent Steering Committee of this series of workshops. Since 2003, he has been the Chair of the SPIE Conference on Image and Signal Processing for Remote Sensing. He has been the founder of the *IEEE Geoscience and Remote Sensing Magazine* for which he has been Editor-in-Chief between 2013–2017. He is currently an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He has been a Distinguished Speaker of the IEEE Geoscience and Remote Sensing Society between 2012 and 2016.