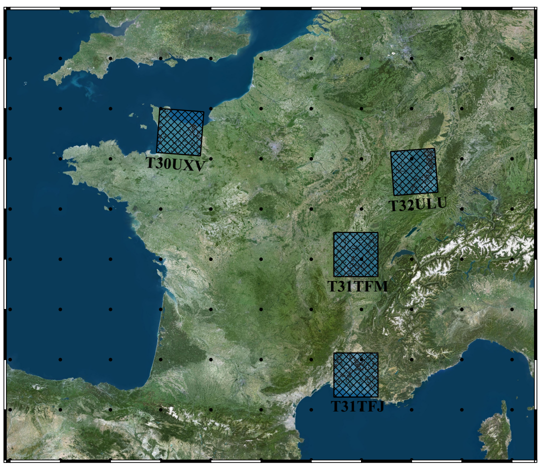
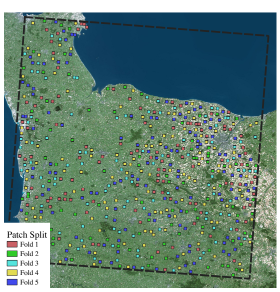
Panoptic Agricultural Satellite image TIme Series (PASTIS) Dataset Documentation

Vivien Sainte Fare Garnot　　Loic Landrieu
LASTIG, Univ. Gustave Eiffel, ENSG, IGN, F-94160 Saint-Mande, France
{vivien.sainte-fare-garnot, loic.landrieu}@ign.fr



(a) Location of the four tiles.



(b) Selected patches.

(c) Single patch.

Figure 1: **Data Location.** Spatial distribution of the four Sentinel tiles used in PASTIS 1a, and of the selected patches of tile T30UXV 1b. We show an example of patch in 1c, and highlight with red circles examples of parcels that are mostly outside of the patch's extent and thus annotated with the void label. The green circle highlight a parcel partially cut off by the patch borders, but with sufficient overlap to be kept as a valid parcel.

The PASTIS dataset is a dataset for evaluation of semantic and panoptic segmentation of agricultural parcels from Satellite Image Time Series (SITS) [4].

Overview. The dataset is composed of 2433 square 128×128 patches with 10 spectral bands and at 10m resolution, obtained from the open-access Sentinel-2 platform. [1] For each patch, we stack all available acquisitions between September 2018 and November 2019, forming our four dimensional multi-spectral SITS: $T \times C \times H \times W$.

The publicly available French Land Parcel Identification System (LPIS) allows us to retrieve the extent and content of all parcels within the patches, as reported by the farmers. Each patch pixel is annotated with a semantic label corresponding to either the parcels' crop type or the background class. The pixels of each unique parcel in the patch receive a corresponding instance label. The French Payment Agency estimates the accuracy of the LPIS annotations as over 98% regarding crop types. While there are no official quantitative assessments regarding parcel surfaces, we performed an extensive visual inspection and failed to observe delineation errors.

Dataset Extent. The SITS of PASTIS are taken from 4 different Sentinel-2 tiles in different regions of the French metropolitan territory as depicted in Figure 1a. These regions cover a wide variety of climates and culture distributions. Sentinel tiles span 100×100km and have a spatial resolution of 10 meter per pixel. Each pixel is characterized by 13 spectral bands. We select all bands except the atmospheric bands B01, B09, and B10. Each of these tiles is subdivided in square patches of size 1.28×1.28km (128×128 pixels at 10m/pixel), for a total of around $24,000$ patches. We then select $2,433$ patches (10% of all available patches, see Figure 1b), favoring patches with rare crop types in order to decrease the otherwise extreme class imbalance of the dataset.

[1] https://scihub.copernicus.eu

Satellite Imagery We use the L2A Sentinel-2 imagery prepared by THEIA. All bands are spatially resampled to a 10m/pixel resolution with bilinear interpolation.

Nomenclature The FLPIS uses a 73 class breakdown for crop types. We select classes with at least 400 parcels and with samples in at least 2 of the 4 Sentinel-2 tiles. This leads us to adopt a 18 classes nomenclature, presented in Figure 2. Parcels belonging to classes not in our 18-classes nomenclature are annotated with the *void* label, see below.

Patch Boundaries. The FLPIS allows us to retrieve the pixel-precise borders of each parcel. We also compute bounding boxes for each parcel. The parcels' extents are cropped along the extent of their 128×128 patch, and the bounding boxes are modified accordingly. Parcels whose surface is more than 50% outside of the patch are annotated with the *void* label, see Figure 1c.

Small parcels To avoid degenerate cases where the size of the parcel is too small compared to the resolution of Sentinel-2, we chose to remove some agricultural parcels from the dataset based on the following geometrical criteria:

- Parcels that have a surface smaller that 800m^2 (*i.e.* 8 Sentinel-2 pixels)
- Parcels for which the ratio of the area over the perimeter is smaller than 10 meters.

Such parcels are annotated with the background label.

Void and Background Labels. Pixels which are not within the extent of any declared parcels are annotated with the background "stuff" label, corresponding to all non-agricultural land uses. For the semantic segmentation task, this label becomes the 20-th class to predict. In the panoptic setting, this label is associated with pixels not within the extent of any predicted parcel. We do not compute the panoptic metrics for the background class, since our focus is on retrieving the parcels' extent rather than an extensive land-cover prediction. In other words, the reported panoptic metrics are the "things" metrics, which already penalize parcels predicted on background pixels by counting them as false positives.

The void class is reserved for *out-of-scope* parcels, either because their crop type is not in our nomenclature or because their overlap with the selected square patch is too small. We remove these parcels from all semantic or panoptic metrics and losses. Predicted parcels which overlap with an IoU superior to 0.5 with a void parcel are not counted as false positive or true positive, but are simply ignored by the metric, as recommended in [1].

Label and Color	Class Name	Number of parcels
0	*Background*	-
1	Meadow	31292
2	Soft winter wheat	8206
3	Corn	13123
4	Winter barley	2766
5	Winter rapeseed	1769
6	Spring barley	908
7	Sunflower	1355
8	Grapevine	10640
9	Beet	871
10	Winter triticale	1208
11	Winter durum wheat	1704
12	Fruits, vegetables, flowers	2619
13	Potatoes	551
14	Leguminous fodder	3174
15	Soybeans	1212
16	Orchard	2998
17	Mixed cereal	848
18	Sorghum	707
19	*Void label*	35924

Figure 2: Color code of our class nomenclature, and the number of parcel per class.

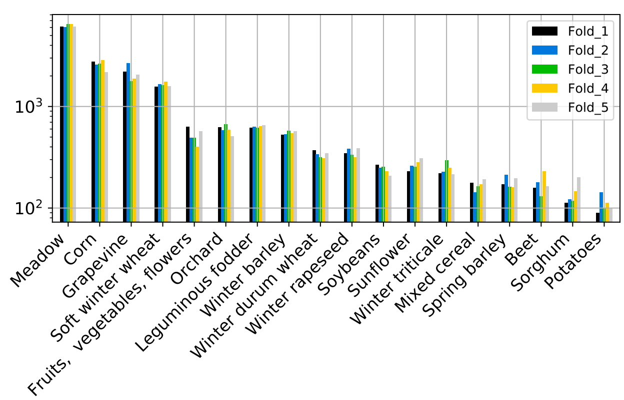


Figure 3: Class distribution for the five folds (in log-scale).

Cross-Validation. The 2,433 selected patches are randomly subdivided into 5 splits, allowing us to perform cross-validation. The official 5-fold cross-validation scheme used for benchmarking is given in Table 1. In order to avoid heterogeneous folds, each fold is constituted of patches taken from all four Sentinel tiles. We also chose folds with comparable class distributions, as measured by their pairwise Kullback-Leiber divergence. We show the resulting class distribution for each fold in Figure 3. Finally, we prevent adjacent patches from being in different folds to avoid data contamination. Geo-referencing metadata of the patches and parcels is included in PASTIS, allowing for the constitution of geographically consistent folds to evaluate spatial generalization.

Fold	Train	Val	Test
I	1-2-3	4	5
II	2-3-4	5	1
III	3-4-5	1	2
IV	4-5-1	2	3
V	5-1-2	3	4

Table 1: Official 5-fold cross validation scheme. Each line gives the repartition of the splits into train, validation and test set for each fold.

Temporal Sampling. The temporal sampling of the sequences in PASTIS is irregular: depending on their location, patches are observed a different number of times and at different intervals. This is a result of both the orbit schedule of Sentinel-2 and the policy of Sentinel data providers not to process tile observations identified as covered by clouds for more than 90% of the tile's surface. As this corresponds to the *real world* setting, we decided to leave the SITS as is, and thus to encourage methods that can favourably address this technical challenge. As a result, the proposed SITS are constituted of 33 to 61 acquisitions.

Clouds Cover. Even after the automatic filtering of predominantly cloudy acquisitions, some patches are still partially or completely obstructed by cloud cover. We opt to not apply further pre-processing or cloud detection, and produce the raw data in PASTIS. Our reasoning is that an adequate algorithm should be able to learn to deal with such acquisitions. Indeed, robustness to cloud-cover has been experimentally demonstrated for deep learning methods by Rußwurm and Körner [2, 3].

Credits The satellite imagery used in PASTIS was retrieved from THEIA: "Value-added data processed by the CNES for the Theia www.theia.land.fr data cluster using Copernicus data. The treatments use algorithms developed by Theia's Scientific Expertise Centres. "

The annotations used in PASTIS stem from the French land parcel identification system produced by IGN, the French mapping agency.

This work was partly supported by ASP, the French Payment Agency.

References

[1] Alexander Kirillov, Ross Girshick, Kaiming He, and Piotr Dollár. Panoptic feature pyramid networks. In *CVPR*, 2019.

[2] Marc Rußwurm and Marco Körner. Convolutional LSTMs for cloud-robust segmentation of remote sensing imagery. *NeurIPS Workshops*, 2018.

[3] Marc Rußwurm and Marco Körner. Self-attention for raw optical satellite time series classification. *ISPRS*, 2020.

[4] Vivien Sainte Fare Garnot and Loic Landrieu. Panoptic segmentation of satellite image time series with convolutional temporal attention networks. *arxiv*, 2021.