DENETHOR: Supplementary Material

A Data Availability

The dataset is already public and used as part of a competition. Up-to-date links for data access and the competition can be found on the GitHub repository. The test set is not yet public to ensure fairness in the competition but will be released after the competition ends in December. We confirm that we bear responsibility in case of violations of rights.

B Additional Results

Table A1: Macro F1 Score of Benchmark Models on DENETHOR test set in the year 2019 by spatial (rows) and temporal encoder (columns). Just like Table 2 in the main paper with a different metric.

Spatial Encoder	Temporal Encoder			
	TempCNN	MSResNet	LSTM	Transformer
ResNet18	30.96%	29.30%	18.44%	14.72%
SqueezeNet	34.73%	25.10%	5.86%	11.19%
MobileNetv3	32.79%	34.17%	19.89%	26.24%
Pixel Average	49.17%	48.64%	23.82%	33.83%
	Pixel-Set Encoding + Self-Attention			
PselTae	58.12%			
PseTae	54.50%			
	Ablation Scores			
PselTae (2018)	71.76%			
PselTae (Val)	83.23%			

C Datasheet for Datasets Documentation

C.1 Motivation For Datasheet Creation

Why was the datasheet created? The dataset was created to address the gap of available daily time series data for crop type mapping in remote sensing. With this, we want to foster methodological development of new algorithms for these kinds of inputs.

Has the dataset been used already? No, the dataset has not been used apart from our benchmark tests.

What (other) tasks could the dataset be used for? As we outline in Section 3, crop or field instance segmentation could be other interesting tasks to study. Beyond that, the dataset provides some opportunities for the study of declouding and superresolution as well.

Submitted to the 35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks. Do not distribute.

Who funded the creation of the dataset? The creation of the dataset was supported by the German Ministry of Commerce under the grant DynamicEarthNet (FKZ 50EE2005). The Fusion imagery was provided by Planet to support the development of new algorithms for crop type mapping in the scientific community.

C.2 Datasheet Composition

What are the instances? The instances are crop fields in Germany that are part of a public registry.

How many instances are there in total (of each type, if appropriate)? There are about 4,500 fields in total

What data does each instance consist of ? Each instance is a multipolygon in a .geojson, a popular vector file format for geodata. For each instance, a field ID and its crop type are given.

Is any information missing from individual instances? No, the state registry lists the crop information for all fields in our area of interest.

Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)? Not applicable since the instances are field boundaries.

Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? Our dataset contains all instances of crop fields in our study area. Naturally, however, this is only a subset of crop fields in the country let alone the whole planet. As the distribution of crop types has large spatial dependencies, our sample would, at best, only be representative of a specific climate.

There are two challenges that make a global sample of crop fields difficult: First, the existence and availability of ground truth in sufficient quality in a geographic region is anything but random. Second, local crop types and their documentation are not straightforward to integrate across regions and countries. This is why crop type classification datasets currently tend to be primarily from one spatial region and ours is no exception. As access to crop data and satellite imagery becomes more available, we hope that a higher geographic diversity will be within reach in the future.

Are there recommended data splits (e.g., training, development/validation, testing)? Yes, we present two tiles with one designated to training/validation and the other one for testing. A geographic split of train and test fields is essential to prevent spatial label leakage.

Are there any errors, sources of noise, or redundancies in the dataset? Not to our knowledge. The information about crop fields and types is considered high-quality data but the possibility of single mistakes there can not be excluded.

Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)? The dataset is self-contained.

C.3 Collection Process

What mechanisms or procedures were used to collect the data The crop data was already collected and is publicly available.

How was the data associated with each instance acquired? The crop types are based on self-reports by farmers which are verified by the government in a series of validity checks.

If the dataset is a sample from a larger set, what was the sampling strategy? Not applicable because we did not sample fields.

Who was involved in the data collection process (e.g., students, crowdworkers, contractors) and how were they compensated (e.g., how much were crowdworkers paid)? Not applicable since the field data was not collected by us.

Over what timeframe was the data collected? 2018-2019

C.4 Data Preprocessing

Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? We remove fields below the size of 1000m² but this can be adjusted in the dataloader. So it is rather a parameter than a cleaning process and is left to the user. We therefore skip the rest of the section as suggested by the questionnaire.

C.5 Dataset Distribution

How will the dataset be distributed? The dataset will be distributed via MediaTUM, the dataset sharing platform of the Technical University of Munich. **When will the dataset be released/first distributed? What license (if any) is it distributed under?**

The dataset will be released in October under a CC-BY license as part of a competition. Evaluations on the test set will be possible via the challenge but the test set itself will only be published after the challenge has concluded in December.

Are there any copyrights on the data? The crop field information belongs to the German Government © GeoBasis-DE/LGB (2018/19), and is available under the GovData 2.0 licence (https://www.govdata.de/dl-de/by-2-0).

Are there any fees or access/export restrictions? Not to our knowledge.

C.6 Dataset Maintenance

Who is supporting/hosting/maintaining the dataset? The data is hosted and maintained by the chair for Data Science in Earth Observation at the Technical University of Munich.

Will the dataset be updated? If so, how often and by whom? It is not our intention to update the data itself. If this was necessary, the primary point of contact is xiaoxiang.zhu@dlr.de We will, however, only release the ground truth data for the test set after the end of the competition in early December. The evaluation of predictions for the test set will be done via a public server during the competition where the reference data remains hidden to the participants.

How will updates be communicated? (e.g., mailing list, GitHub) On the DENETHOR Github Repository

If the dataset becomes obsolete how will this be communicated? On the DENETHOR Github Repository

Is there a repository to link to any/all papers/systems that use this dataset? There are no papers that use this dataset yet but we may provide this in the future.

If others want to extend/augment/build on this dataset, is there a mechanism for them to do so? There is no designated mechanism but our dataset could be augmented with additional input data sources or a larger geographic coverage.

C.7 Legal and Ethical Considerations

Were any ethical review processes conducted (e.g., by an institutional review board)? No.

Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctorpatient confidentiality, data that includes the content of individuals non-public communications)? No the data only contains satellite images and crop field information.

Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? No.

Does the dataset relate to people? No, skipping the rest of the section as suggested by the questionnaire.