



## A Scalable Deep Learning Pipeline for Mapping Forest Disturbances

---

Loïc Dutrieux, Keith A. Araño, Nelly Gaillard and Pieter S.A. Beck

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

October 12, 2023

# A SCALABLE DEEP LEARNING PIPELINE FOR MAPPING FOREST DISTURBANCES

*Loïc Dutrieux<sup>1</sup>, Keith A. Araño<sup>\*2</sup>, Nelly Gaillard<sup>1</sup>, Pieter S.A. Beck<sup>1</sup>*

<sup>1</sup>European Commission, Joint Research Centre, Italy

<sup>2</sup>Arcadia Sistemi Informativi Territoriali SRL, Italy

## ABSTRACT

We present a scalable and flexible approach for mapping forest disturbances over large areas. The approach harnesses the power of a spatio-temporal deep learning model, enabling it to capitalize on complex patterns while demanding minimal data pre-processing efforts. Our methodology enables rapid map production, offering a streamlined workflow. The approach was demonstrated over central Europe, covering an area of approximately 900,000 km<sup>2</sup>. Utilizing a modest cluster, the processing time amounted to just 36 hours. The results, produced at 20 m spatial resolution, exhibit coherent patterns and promising accuracy values, with an overall accuracy of 96%. We classify disturbances into four categories, marking a significant stride towards forest disturbance attribution. Identifying areas for improvement, we aim to reduce residual artifacts and enhance accuracy by incorporating higher quality training data. Future work will focus on refining the model architecture and expanding the dataset coverage to further optimize the approach’s performance and accuracy.

**Index Terms**— Sentinel-2, time-series, forest, deep-learning

## 1. INTRODUCTION

Forests, covering around 35% of Europe, provide essential ecosystem services including biodiversity conservation and climate regulation [1]. Their significant role in our planet’s ecological balance necessitates accurate, frequent, and timely monitoring to inform decisions of policy makers and forest managers.

Various methodologies for mapping forest disturbances have been proposed, with some operational at large scales [2, 3]. These approaches typically involve some form of image selection, whether manually or algorithmically driven, which can limit the input data used and pose obstacles to efficient and timely map production. A handful of truly multi-temporal approaches have been proposed, yet, to our knowledge, their scalability remains unproven and has only been tested on a limited scale [4, 5]. It is important to note that temporal segmentation approaches like LandTrendr, despite being inherently multi-temporal, are more suited to extracting

temporal trajectories and breakpoints, making them excellent for retrospective analyses over long periods, but less effective for detecting recent changes [6].

In this paper, we explore the potential of the underutilized Sentinel-2 satellite data, which offers high-resolution, multi-spectral data at a rapid frequency. Challenges such as heterogeneous temporal sampling due to cloud cover and residual clouds, despite default cloud masking, pose difficulties in transferring conventional methods and require algorithms that can robustly handle variable length and density time-series.

Our exploration leads us towards a deep learning approach to harness the potential of Sentinel-2 data for a more flexible and scalable disturbance mapping method. The deep learning model operates on raw satellite data, providing a novel approach in this domain.

The rest of the article is structured as follows: We first outline our strategy for generating training data, followed by a brief overview of the model architecture. Subsequently, we report and discuss performance metrics and the results of a scalability experiment. This exploration, while not claiming to fill the gap completely, contributes to the ongoing efforts in enhancing forest disturbance mapping at scale and in a timely manner.

## 2. MATERIALS AND METHODS

To demonstrate the flexibility and scalability of the proposed approach, the study focuses on 7 countries of Central Europe (Austria, Czech Republic, Germany, Hungary, Poland, Slovakia and Slovenia). With an area of nearly 900,000 km<sup>2</sup> and a broad range of environmental conditions and forest disturbance types, this region offers great opportunities and challenges for demonstration and testing. The following section describes (1) the iterative, semi-supervised, approach employed to obtain training data, (2) the architecture of the spatio-temporal model we used, (3) some elements of model training and (4) the scalable and flexible map prediction strategy we utilized.

### 2.1. Training data generation

To generate a labeled dataset of forest dynamics, we applied a semi-automated protocol over 15 40 by 40 km areas of

<sup>\*</sup>Consultant for the Joint Research Center

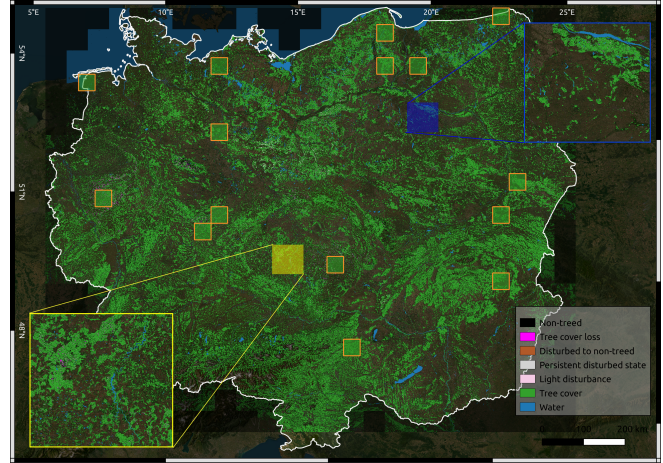
interest (AOI) spread across Central Europe (see Fig. 1 for the footprints of the AOIs). Although the resulting data remains imperfect, the approach allows to efficiently create an adequate amount of annotated data required for subsequent model training. For each AOI independently, the protocol consists of the following steps:

1. Image selection: Manual selection of three cloud-free summer images for the years 2019, 2020 and 2021.
2. Identification and labeling of initial examples: Initial examples of various land trajectories, such as persistent non-treed areas, persistent treed areas, and tree cover loss, are identified and digitized within the selected images. These examples serve as the starting point for training a simple tri-temporal pixel-based classifier.
3. Model training and prediction: Based on digitized training samples, a random forest classifier is used to classify each pixel of the AOI into one of the land trajectories.
4. Result inspection: The quality of the produced land trajectories layer is assessed by visual inspection. By confronting predicted patterns with various sources of information (input Sentinel-2 images, very high resolution layers, hypertemporal Planet data, model uncertainty), the operators can identify mislabeled pixels.
5. Enrichment of Training Data: Training data is enriched, particularly for areas and patterns presenting high model confusion.
6. Iteration: The operators iteratively repeats steps 3 to 5 until they consider that no significant improvement can be achieved.

The annotated dataset resulting from applying this protocol over the 15 AOIs covers an area of 24,000 km<sup>2</sup> for two one-year time periods, and divides the land dynamics into 7 classes (persistent non-treed, persistent tree cover, tree cover loss, light disturbance, persistent disturbed state, disturbed to non-treed, water).

## 2.2. Model architecture

To address the challenge of scalable and detailed mapping of forest dynamics, we used a spatio-temporal deep learning architecture. Deep learning presents an attractive option due to (i) its ability to automatically extract meaningful features from multi-dimensional data, and (ii) its capacity to internally handle data heterogeneity such as noise and irregular temporal sampling. In the context of forest mapping, this means that the deep learning model can capture important contextual information and associate complex spectro-spatio-temporal signatures to forest dynamic patterns which is crucial for accurate mapping. The handling of heterogeneous data, which



**Fig. 1:** Map of forest dynamics between July 2020 and July 2021. The green squares with orange outline are the footprints of the AOIs used for training data generation produced following the protocol described in 2.1

often poses limitations for classical machine learning methods, is a crucial characteristics for flexibility and scalability allowing to take full advantage of the entire stack of temporal information contained in Sentinel-2 data.

Following these considerations, we adapted a deep-learning architecture initially designed for volumetric segmentation named UNet3D [7]. This model extends the classical UNet by incorporating 3D convolutional filters, enabling the simultaneous consideration of spatial context and dependencies in an additional dimension, in this case, time.

## 2.3. Model training

Both the training and prediction phases of our model were conducted using 20 m Sentinel-2 data that exhibited less than 30% cloud cover. We specifically utilized the spectral bands B2, B3, B4, B5, B6, B7, B8A, B11, and B12, and complemented them with three vegetation indices: Normalized Difference Vegetation Index (NDVI), Normalized Burn Ratio (NBR), and Normalized Difference Moisture Index (NDMI). Although the model should allow working on unfiltered data, we filtered clouds and shadows using the Scene Classification Layer (SCL) provided with level 2A Sentinel-2 data. To establish a consistent temporal framework for training, we re-sampled the data into a zero-filled, 10-day interval vector, resulting in a total of 46 time steps. This temporal resampling approach allows to maintain temporal coherence and align the data across different timestamps, enabling effective temporal analysis. The training process was performed end to end, leveraging the capabilities of an NVIDIA Tesla V100 32GB GPU.

## 2.4. Map prediction

For the map prediction step, which consists in running the deep learning model in inference mode, CPU-based computation was utilized. While individual CPUs generally have lower performance compared to GPUs, their higher availability makes it easier to horizontally scale the process, accommodating larger workloads. The computing load was distributed on a Dask cluster, allowing for efficient utilization of computational resources and parallelized execution of tasks, enabling faster overall map production [8].

All the Sentinel-2 data required for the map prediction are available on the EOS distributed file system (EOS) of the Big Data Analytic Platform (BDAP) of the Joint Research Center (JRC) and indexed in a SpatioTemporal Asset Catalog (STAC) [9, 10]. We used the `pystac-client` python library to query the STAC catalog and retrieved the data as a lazy data cube thanks to the `odc-stac` package [11]. Some data preparation steps, as described in section 2.3, were necessary and performed on the fly as part of the prediction pipeline.

The desired output map was set to be in the `EPSG:3035` coordinate reference system (CRS) and two strategies were considered to achieve this. The first strategy involves using MGRS (Military Grid Reference System) tiles for task definitions and processing each tile in its respective input CRS, typically one of the UTM zone intersecting Central Europe. Once every MGRS tile has been predicted, the individual raster files need to be mosaicked, handling different CRS and overlap between tiles as a post-processing step. The second strategy involves defining a regular grid in the desired coordinate reference system (`EPSG:3035`) and warping the input data to align with that CRS. Using that second strategy, the result of the predictions are directly produced in the target CRS and no further post-processing steps are required. This resembles a processing pipeline from an Analysis Ready Data (ARD) collection, except that the harmonized input data is generated on the fly, from the raw collection. While that second strategy implies higher computational requirements, increasing the processing time for map production, the absence of post-processing steps reduces the amount of mosaicking artifacts and makes the overall approach more agile; it was therefore preferred for the current demonstration.

## 3. RESULTS AND DISCUSSION

### 3.1. Map production and qualitative assessment

The production of the final map using the approach described in section 2.4 required 36 hours of processing on a single machine cluster spread over 38 cores. This is a very encouraging result given the size of the study area and suggests a high potential to scale a similar approach to much larger areas, and obtain forest disturbance layers covering the whole of Europe within days. As can be seen in Fig. 1, the

resulting map covers the entire study area without interruptions and very few processing artifacts (e.g. visible tile edges) could be found. A detailed visual inspection of the result revealed generally coherent patterns of various landscape elements including water bodies, tree cover areas and forest disturbances. Well known hot-spots of forest disturbances due to widespread bark beetle outbreaks in Germany, Czechia and Slovakia are clearly visible and include all four disturbance classes. Unsurprisingly, the highest confusion in classification appeared to be found between subtle disturbance classes, such as persistent disturbed state and transitions from healthy to lightly disturbed forests. This is expected due to the nuanced structural and chemical processes at stake, compared to scenarios like a complete stand removal. We also noticed that North-facing slopes in steep terrain were often classified non-treed, regardless of whether trees were present or not; this can be explained by their low sun illumination combined with the fact the labelled dataset did not present such terrain configuration.

### 3.2. Model performance metrics

By confronting the model predictions to a held out partition of the labelled dataset, we could generate some model performance metrics (Table 1).

**Table 1:** Classification Performance

Class	Producer's Accuracy	User's Accuracy	F1 Score
Persistent Non-Treed	.98	.97	.97
Persistent Tree Cover	.94	.96	.95
Tree Cover Loss	.77	.73	.75
Light Disturbance	.46	.45	.46
Persistent Disturbed State	.34	.27	.30
Disturbed to Non-Treed	.27	.65	.38
Water	.95	.99	.97

Overall Accuracy: .96

The accuracy for the "stable" classes (Persistent Non-Treed, Persistent Tree Cover, Water) is consistently high, as anticipated due to the distinct spectral and temporal signatures associated with these classes. In contrast, notable confusion is observed among the four closely related disturbance classes, indicating the need for further refinement in distinguishing them.

It is important to note that these benchmark metrics, while useful and convenient to get a rapid snapshot of a model's potential, do not necessarily translate into an unbiased estimate of the accuracy of the produced map. Reasons are that:

- The data used for computing these metrics originates from the same data generation protocol described in section 2.1, and while the quality is generally high, it likely still contains errors.
- Accuracy tends to vary in space and time, even when the exact same approach is strictly replicated over a different area or time periods [12]. Map accuracy is there-

fore not an intrinsic characteristic of the model used, but rather a complex context dependent parameter.

- Model benchmark metrics only account for one step of the map production process, the prediction. Other steps, such as tiling strategy that may have an impact on the accuracy of the final map are not captured by these metrics.

A proper accuracy assessment of the produced map will require an independent set of visually interpreted samples obtained through probability sampling [13].

### 3.3. Operationalization potential

The operationalization potential of the approach is noteworthy. It requires minimal pre-processing, relying on a cluster, data archive, and STAC catalogue for producing scalable maps. Despite the utilization of imperfect training data, the approach delivers visually coherent results and demonstrates encouraging model performance metrics. Remaining issues, as described earlier, necessitate further research. Improvements are required in both the model architecture and training data quality, particularly for accurately characterizing complex forest disturbances such as light disturbances and standing dead trees. Nevertheless, the approach partially distinguishes four types of forest disturbances, marking a significant achievement in forest disturbance attribution. Enhancements should focus on addressing identified issues and refining the model architecture and training data quality to increase accuracy and expand the applicability of the approach.

## 4. CONCLUSIONS

In conclusion, we successfully developed a scalable forest disturbance mapping approach by harnessing the capabilities of deep learning. This allows us to generate large area maps with minimal pre-processing requirements. The results obtained from the demonstration in Central Europe are highly promising, showcasing the effectiveness of the approach in capturing and characterizing forest disturbances. However, the study has also identified certain shortcomings and areas that require improvement. These insights will guide future work, enabling us to address these limitations and enhance the approach's performance and accuracy. With ongoing research and refinement, we are confident that our approach will make a substantial contribution to enhancing operational capacity in achieving detailed mapping of forest disturbances in temperate environments.

## REFERENCES

[1] Forest Europe. *State of Europe's forests 2020*. Ministerial Conference on the Protection of Forests in Europe, Madrid, 2020.

- [2] Matthew C Hansen, Peter V Potapov, Rebecca Moore, et al. High-resolution global maps of 21st-century forest cover change. *science*, 342(6160):850–853, 2013.
- [3] Collin Homer, Jon Dewitz, Limin Yang, et al. Completion of the 2011 national land cover database for the conterminous united states—representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5):345–354, 2015.
- [4] Roberto Interdonato, Dino Ienco, Raffaele Gaetano, and Kenji Ose. Duplo: A dual view point deep learning architecture for time series classification. *ISPRS journal of photogrammetry and remote sensing*, 149:91–104, 2019.
- [5] Marc Rußwurm and Marco Körner. Multi-temporal land cover classification with sequential recurrent encoders. *ISPRS International Journal of Geo-Information*, 7(4):129, 2018.
- [6] Robert E Kennedy, Zhiqiang Yang, and Warren B Cohen. Detecting trends in forest disturbance and recovery using yearly landsat time series: 1. landtrendr—temporal segmentation algorithms. *Remote Sensing of Environment*, 114(12):2897–2910, 2010.
- [7] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, et al. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, pages 424–432. Springer, 2016.
- [8] Dask Development Team. *Dask: Library for dynamic task scheduling*, 2016. URL <https://dask.org>.
- [9] P. Soille, A. Burger, D. De Marchi, et al. A versatile data-intensive computing platform for information retrieval from big geospatial data. *Future Generation Computer Systems*, 81: 30–40, 2018. ISSN 0167-739X. doi: [10.1016/j.future.2017.11.007](https://doi.org/10.1016/j.future.2017.11.007).
- [10] Matthew Hanson. The open-source software ecosystem for leveraging public datasets in spatio-temporal asset catalogs (stac). In *AGU Fall Meeting Abstracts*, volume 2019, pages IN23B–07, 2019.
- [11] Brian Killough. Overview of the open data cube initiative. In *IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium*, pages 8629–8632. IEEE, 2018.
- [12] Katsuto Shimizu, Tetsuji Ota, and Nobuya Mizoue. Accuracy assessments of local and global forest change data to estimate annual disturbances in temperate forests. *Remote Sensing*, 12(15):2438, 2020.
- [13] Pontus Olofsson, Giles M Foody, Martin Herold, Stephen V Stehman, et al. Good practices for estimating area and assessing accuracy of land change. *Remote sensing of Environment*, 148:42–57, 2014.