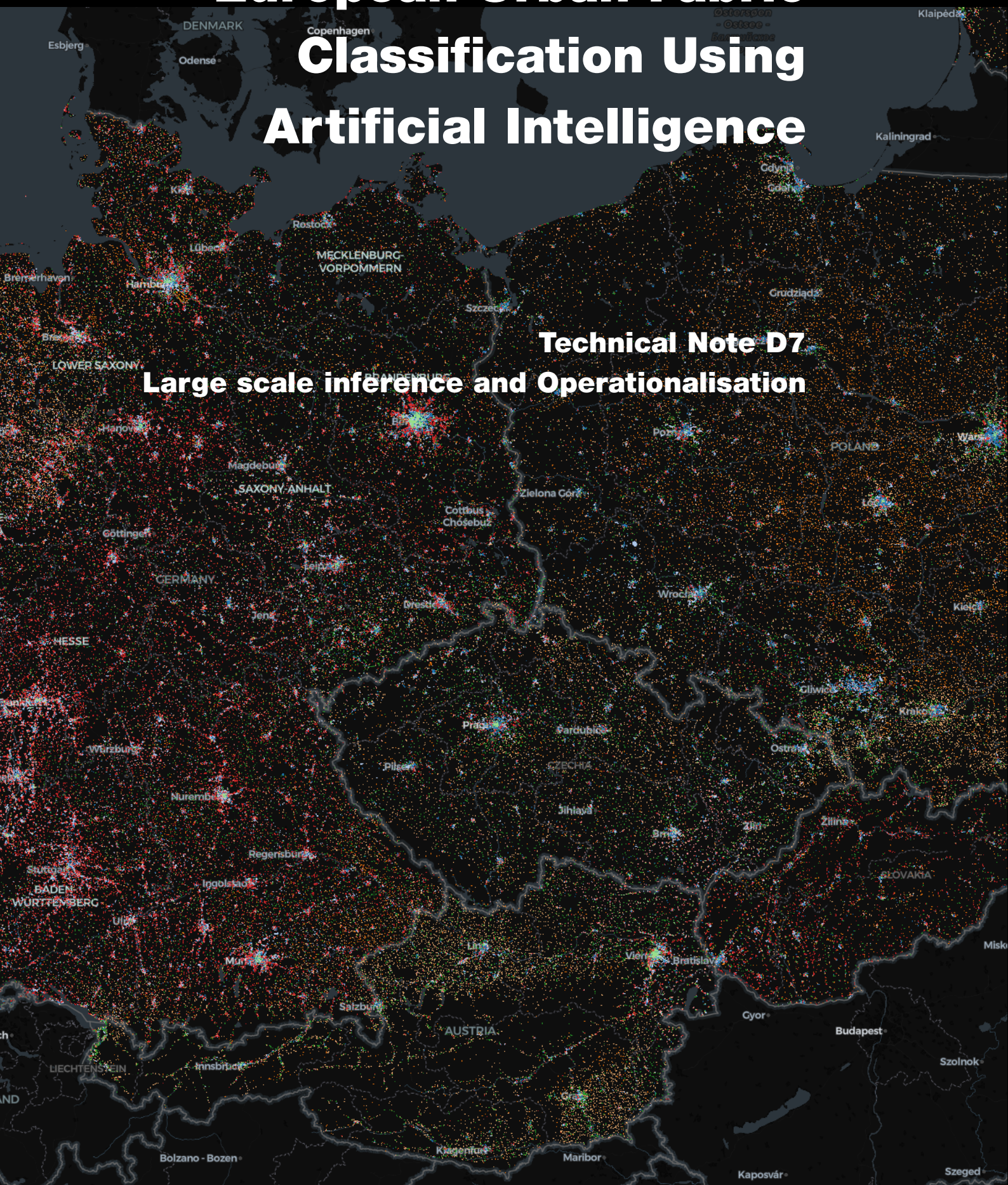


European Urban Fabric Classification Using Artificial Intelligence

Technical Note D7
Large scale inference and Operationalisation



Large scale inference and Operationalisation

Technical Note D7

Krasen Samardzhiev¹, Barbara Metzler², Martin Fleischmann^{1,*} and Daniel Arribas-Bel²

1. Faculty of Science, Charles University, Social Geography and Regional Development, Albertov 6, 128 00, Praha, Czechia
2. The Alan Turing Institute, Science of Cities and Regions, British Library, 96 Euston Road, London NW1 2DB, UK

* *Correspondence:* Martin Fleischmann martin.fleischmann@natur.cuni.cz



The activities reported in this document were carried under a programme of, and funded by, the European Space Agency. View expressed in the publication can in no way be taken to reflect the official opinion of the European Space Agency.

eurofab.org

Table of contents

1	Introduction	1
2	European Morphometric Classification Strategy	2
2.1	Data-readiness	2
2.2	Risk assessment	2
2.3	Europe-wide classification strategy	2
3	European Space-Time Urban Fabric Strategy	3
3.1	Infrastructural, technical and methodological aspects involved in scaling .	3
3.2	Risk assessment	4
3.3	Europe-wide deployment strategy	4
4	Scaled up stakeholder engagement	6
5	Productionisation	8

1 Introduction

This Technical Note presents a strategic roadmap for advancing the EuroFab project’s urban fabric classification system from its current capabilities to a pan-European operational capacity.

Ultimately, EuroFab seeks to pave the road for a world where stakeholders, from local authorities to supranational organisations, are able to track and monitor the pattern of urban development across time, in detail directly relevant for planning, and at scale. The pilot study in Central Europe and the UK consisted of three related parts - two machine learning and AI data processing pipelines, and a stakeholder engagement component. The current pilot study demonstrated the viability of using alternative data sources to produce multiscale descriptions of the built environment, based on analysis of individual buildings, street segments and their interactions. Furthermore, the stakeholder engagement part highlighted the multiple applications which stakeholders see for such a dataset.

The overall strategy can be summarised as consisting of three steps. First, using the first pipeline to generate detailed training data, at a large scale. Second, training the AI vision model on the outputs from step one, to fill data gaps and produce a temporal dimension for the European urban fabric classification. Third, to productise the results and expand the stakeholder engagement in order to drive adoption and derive secondary indicators and datasets for specific applications.

Transitioning the EuroFab system from a successful research project with functional processing pipelines to a robust, reliable, and continuously updated data product presents considerable technical and organizational challenges. Serving a diverse range of stakeholders across Europe with consistent and high-quality urban fabric information requires improvement and upscaling of data processing capabilities, the establishment of a sustainable operational framework that covers potential risks, and a commitment to continuous improvement and user engagement.

The roadmap in this technical note addresses these multifaceted challenges, providing a structured approach to achieve full operationalization.

2 European Morphometric Classification Strategy

2.1 Data-readiness

summary of approach.

Applying the methodology to the full european continent needs to map data sources.

2.2 Risk assessment

Heterogenous data sources. - already know it works, but need to handle exceptions and topological issues and possibly fix the pipeline. Works on overture - easy to use implemented.

Complete data unavailability. - vision model will cover

Unknown urban fabric classes. - this is the work...

2.3 Europe-wide classification strategy

3 European Space-Time Urban Fabric Strategy

3.1 Infrastructural, technical and methodological aspects involved in scaling

The core of EuroFab’s current Earth Observation analytical capability is the Spatial Signatures Pipeline, detailed in the [eurofab-project/eo](#) GitHub repository. This pipeline is engineered to generate spatial signature predictions using satellite imagery as its primary input. First, yt integrates computer vision models to create feature embeddings from satellite imagery. These embeddings are subsequently processed by an XGBoost model, which performs the classification of individual image chips into predefined urban fabric classes. The following aspects of this baseline processing chain need to be adapted and scaled:

- **Sourcing and processing historical European Sentinel 2 data:** The current pipeline is processing historical Sentinel 2 data only for the UK. The data processing and storage have to be expanded to cover the whole of Europe.
- **Classification Labels:** The current XGBoost model assigns each grid cell to one of twelve predefined spatial classes and their aggregations. These classes include: ‘Accessible suburbia’, ‘Connected residential neighbourhoods’, ‘Countryside agriculture’, ‘Dense residential neighbourhoods’, ‘Dense urban neighbourhoods’, ‘Disconnected suburbia’, ‘Gridded residential quarters’, ‘Open sprawl’, ‘Urban buffer’, ‘Urbanity’, ‘Warehouse/Park land’, and ‘Wild countryside’. This typology provides a foundational characterization of diverse urban and peri-urban environments. However, spatial signatures are restricted to the extent of Great Britain. The future development will therefore use a combination of the ground truth data (existing morphological classification) and the predicted classes from the morphological pipeline, where the cadastral data is not available. This change will generate orders of magnitude more training data for the vision and XGBoost models, from all across Europe. This will not just increase the scale of the results, but also their accuracy as shown by the conclusions of the pilot study.
- **Infrastructural Capacity:** The current inference pipeline and its training is not heavily computationally demanding as it was trained on a single high-end GPU. However, the training times were not optimal and for a scaled-up pipeline,

the computational demands will grow significantly. Therefore, it is expected that training and inference will require a small cluster of high-end GPUs, with the inference ran on a yearly basis.

3.2 Risk assessment

The following are the three challenges that need to be tackled for the scaling of the AI models:

- **Large scale evaluation of urban fabric change across time:** The pilot analysis demonstrated that urban fabric classifications exhibit distinct temporal and spatial dynamics, reflecting varying levels of stability and diversity over time. Urban and sub-urban classes displayed higher probabilities of transition, indicating active urban transformation. Conversely, rural classes showed significant stability. As highlighted in the pilot’s results these observed changes may be more related to classifier uncertainty than actual environmental changes.
- **Generalisability Testing:** Evaluating the generalisability of the AI methodological framework is crucial for its wider applicability. Future research should test this modelling approach in different European regions, assessing whether the chosen methods, including data preprocessing, augmentation strategies, spatial embeddings, and classifier architectures, perform consistently. Understanding these factors will help identify potential adjustments needed to ensure reliable predictions when extending the model.
- **Handling Misclassifications:** Misclassifications typically occur between visually similar urban fabric classes, indicating inherent uncertainty in predictions. Incorporating prediction probabilities into a secondary model could help address this issue. By explicitly using probability scores from the initial classification as input for a refinement model — as previously shown by Fleischmann and Arribas-Bel^[3] — we could better distinguish between ambiguous cases. This approach may “smooth” predictions, reducing noise and improving overall classification accuracy. Future work should explore how prediction confidence scores can be systematically utilised, either by employing spatial smoothing algorithms or by applying secondary machine learning models trained specifically to correct uncertain predictions.

3.3 Europe-wide deployment strategy

- **Leveraging Copernicus Services:** The Copernicus programme, with its commitment to free, full, and open data access, is the cornerstone for sourcing satellite imagery. Primarily, data from the Sentinel-2 (multispectral optical) mission will

provide consistent, systematic, and frequently updated coverage of the entire European continent. These datasets are fundamental inputs for the raster-based analysis performed by the EuroFab pipeline.

- **Leveraging the new morphometric classification results as ground truth:** It is expected that changing the ground truth classes from the spatial signatures to the new morphometric results will not lead to changes in the pipeline. Although the spatial signatures capture functional information, almost all morphometric characters used in deriving the new classification are already used in the spatial signatures derivation. Therefore, the current AI vision model pipeline is already taking urban form into account.
- **Generalisability Testing:** The performance of the model will be tested using the same robust generalisability testing framework used in the pilot study's morphological pipeline. This was not possible for the current AI pipeline due to data shortage, however with the expansion of the study area and available data more complex testing scenarios become viable.
- **Evaluation of urban predictions across time** Similarly, the larger quantity of data, as well as the new testing frameworks will enable better model performance and therefore better certainty in the temporal classification.
- **Methodological processing** The current pipeline needs localised information to generate the embeddings and predictions, therefore it is expected that the memory and computing requirements will scale linearly and will not require changes to the underlying methodology. Most of the development time will be spent on pre-processing the data and for training generating the embeddings from the satellite images.

4 Scaled up stakeholder engagement

The stakeholder engagement in the pilot study confirmed the need for a detailed urban classification at a granular level and resulted in multiple feature requests for secondary indicators, specific functionality and reports. A pan-european detailed urban classification can act as the foundation for multiple derived data products adapted to particular use cases. Requested examples included comparisons between countries at different scales - neighbourhoods, cities and regions; access to the building-level morphological data of the taxonomy; and the ability to figure out where newly-designed areas would fall within the taxonomy before they are built. Furthermore, there were multiple requests to combine the classification with other data sources.

The larger scale of the new results will enable us to continue and enhance the stakeholder work:

- **Continue workshops and conference work:** The most successful collaborations and insights came from workshops and conferences. All of these benefits still count for the pan-european Study.
- **Focus on specific European cities and regions:** Stakeholders were typically interested in their local areas or countries, as well as the larger scale data. Even the pilot study extent of five central european countries and the UK proved limiting. For example, the Lithuanian department was interested and we had to adapt and include them in the research. This needs to be scaled up with the new pan-european extent.
- **Collaborations & codevelopment:** The focus of the work was on establishing relationships, and extended collaborations on specific applications and data products was not possible due to the short duration of the project. True co-development in an operational context means evolving beyond simple feedback collection. It involves establishing formal partnerships where stakeholders actively contribute resources – such as local datasets for training/validation, domain expertise, or personnel for joint validation campaigns – and, in return, share in the benefits and responsibilities of the service. This could involve Memoranda of Understanding with national mapping agencies, urban planning institutes, or research consortia, thereby deeply embedding the EuroFab service into their operational workflows and decision-making processes.
- **Developing derived products:** For specific strategic use cases we can extend the stakeholder work and guide the development of derived products. This work will

4 Scaled up stakeholder engagement

also result in detailed examples and documentation of how to use the data products for the development of derived indicators, for different policy areas. Some applications that came up during the development work were combining the classification with land use data and geomdemographic factors, generating summaries of the morphometric character of cities, access to the full taxonomy, as well as raw morphometric characters data at the building level.

- **Relationships and engagement with people developing derived products:** There were multiple proposals for companies to develop their own derived products for specific cases. These include local and private companies in the urban planning and environmental management sectors. Ongoing engagement and support can be provided for them.
- **Open source community:** EuroFab’s commitment to openness is a key asset for its long-term impact and sustainability. The project’s commitment to “open-source software, algorithms and open datasets” [2] must be continued and expanded. This includes maintaining public GitHub repositories (e.g., `eurofab-project/morphometrics`, `eurofab-project/eo`) for all relevant code, tools, and documentation. The work can also be extended to establish and maintain online forums, mailing lists, and developer channels (e.g., Slack, Discord) to provide support to users and foster a collaborative environment for contributors.

5 Productionisation

Moving EuroFab from a research project to an operational service requires a well-defined framework covering deployment, service delivery, quality assurance, and iterative development based on user needs.

- **Defining Operational Service Levels and Output Products:** A tiered approach to service delivery can cater to diverse user needs and ensure sustainability.
 - *Baseline Products:* These would include pan-European urban fabric maps based on the core classification scheme, updated regularly (e.g., annually or biennially). These products should be openly accessible.
 - *Derived Indicators:* Offer services to generate higher-level indicators from the base urban fabric data. Examples include metrics related to green space accessibility, urban sprawl patterns, impervious surface percentage, building density, or indicators supporting SDG11 monitoring. Many examples of urban indicators exist which could be adapted or inspired by, such as those related to sustainable urban development, quality of life, or environmental performance.
- **Potential Integration with Existing European EO Platforms:** To maximize impact and avoid creating isolated systems, EuroFab services should be integrated with or exposed through relevant existing European platforms.
- **Model Retraining and Update Cadence:** - A clear strategy for model retraining is needed. Triggers for retraining could be yearly updates or the new availability of significant new volumes of high-quality training data.
- **User Adoption and Trust:** Implement transparent and publicly accessible Cal/Val procedures and results; integrate XAI techniques to make model outputs understandable; invest heavily in stakeholder engagement, co-design processes ³, training, and clear documentation.