Scalable Forest Biodiversity Mapping from Aerial Imagery under Extreme Label Scarcity

Ivana Pallister[1], Sriram Swaminathan[2], Dr. Joep Burger[1], Marko Roos[1], Shaya van Houdt[1], Dr. Patrick Bogaart[1]
1) Statistics Netherlands (CBS), Heerlen, the Netherlands
2) Maastricht University, Maastricht, the Netherlands

MOTIVATION

- EU Nature Restoration Regulation legally requires forest monitoring
- Dutch forest inventories provide ~800 noisy field plots/year, orders of magnitude fewer than deep learning needs
- Aerial wall-to-wall RGB-NIR mosaics available annually at 25cm resolution across all Dutch forests (2016-2025)

How can we most effectively leverage limited labeled data to generate accurate, nationally representative forest maps?

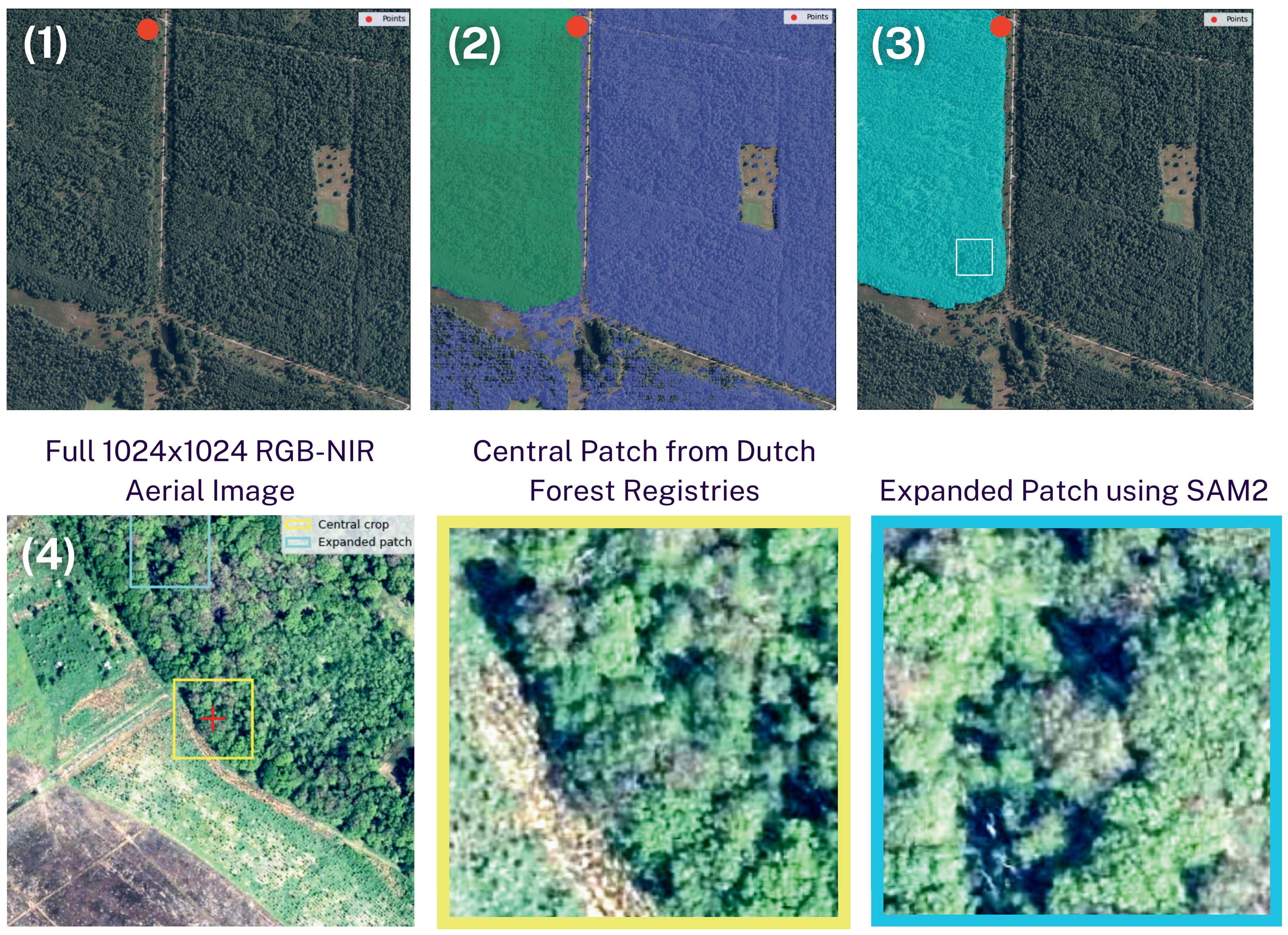
METHODOLOGY

1) Generate pseudo-labeled patches via semantic segmentation

(1) Use Dutch forest registry plot coordinate as point prompt
(2) SAM2 segmentation generates masks for different forest sections
(3) Determine forest section of reference point and preprocess corresponding mask
(4) Sample patches from forest mask and combine with reference label

Aerial Image with Label Reference in Red | Candidate Segmentation Masks | Morphologically Smoothed Agreement Mask

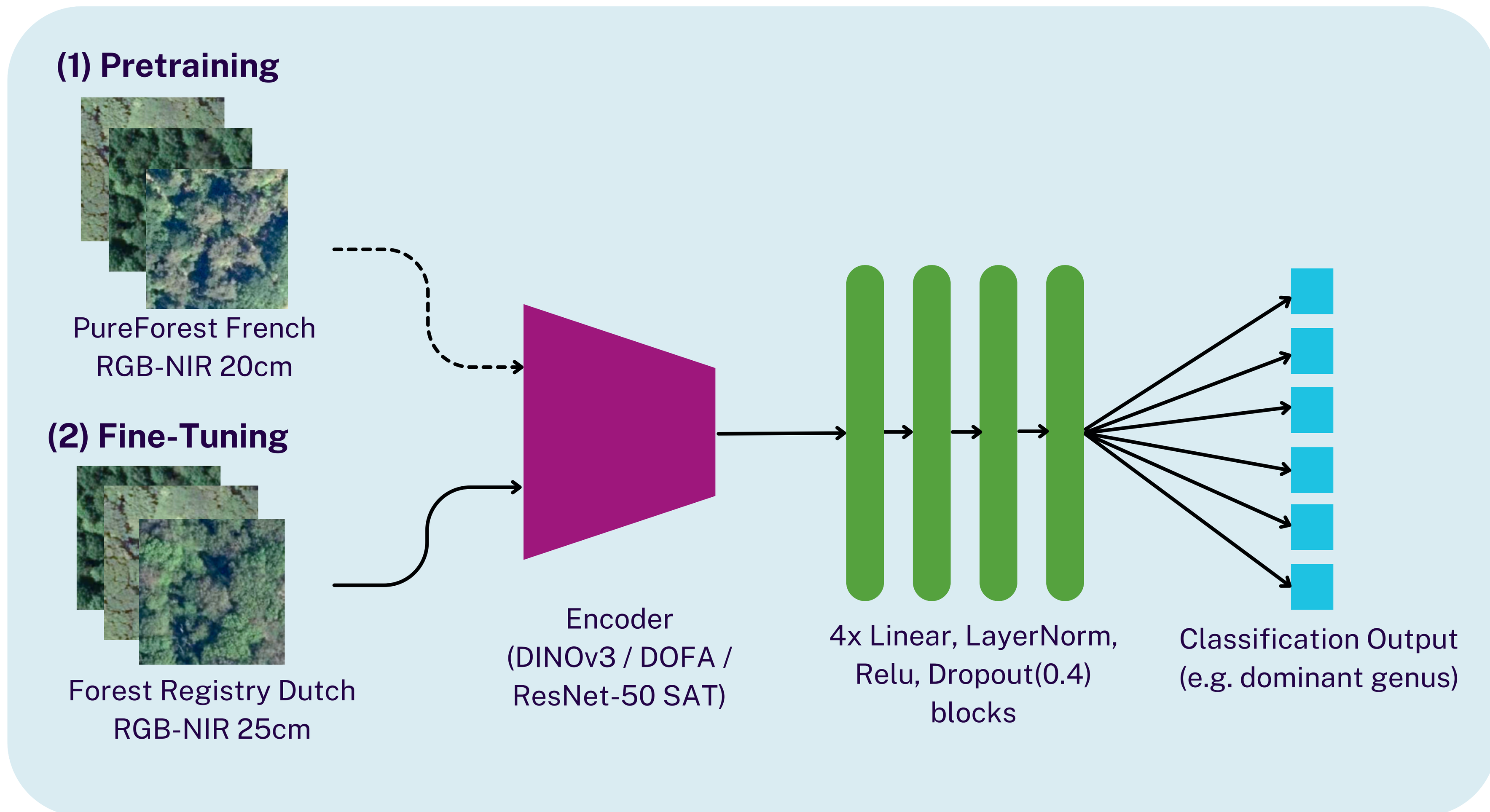
Full 1024x1024 RGB-NIR Aerial Image | Central Patch from Dutch Forest Registries | Expanded Patch using SAM2



2) Transfer learning from other regional aerial image datasets

(1) Pretrain on PureForest dataset
- 50k aerial images of French monospecific forests with 20 cm resolution, annotated with single dominant tree species (13 classes)
- Employ generally pretrained geospatial foundation models as encoder

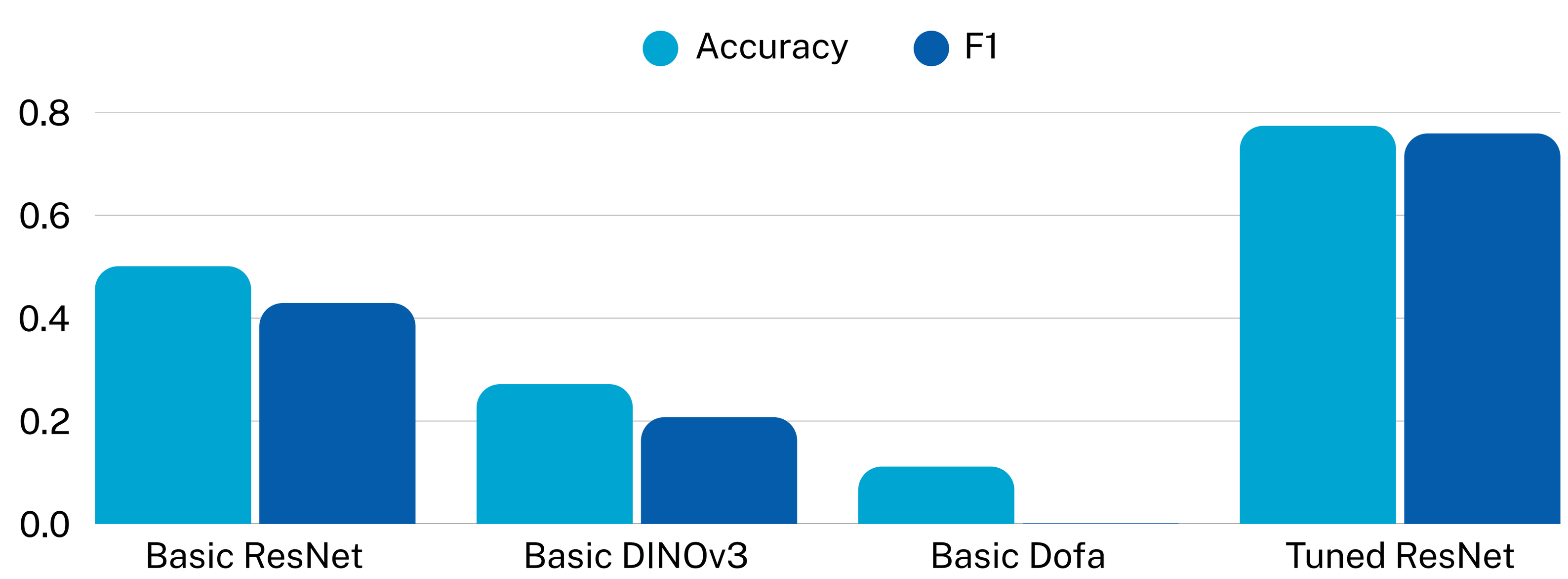
(2) Fine-Tune on limited labeled Dutch data
- ~3500 unique labeled LMF plots per year, heterogeneous stand composition, and systematic geolocation bias at stand boundaries



ONGOING WORK AND CHALLENGES

- Current evaluation on dominant genus classification
  → A necessary first step toward full biodiversity mapping

Proof of Concept:
Training is successful on clean PureForest data, metrics on 20% test set:



Additional challenges with Dutch forest registry data

- More diverse forests compared to French monospecific forests
- Labels were designed for traditional statistics:

Plot coordinate at forest edge | Label present, forest absent | Severe geolocation offset

NEXT STEPS

Data Quality Pipeline
Coverage threshold filter, NVDI label validation, GPS uncertainty filter

Iterative Label Refinement
Expanded patch quality depends on the original noisy labels, making this inherently circular. Exploring ways to break that cycle, using model predictions to flag and correct low-confidence label-patch pairs.

RELATED LITERATURE

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
Ravi, N., Kirillov, A., Rolland, C., Gustafson, L., Mao, H., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2024). SAM 2: Segment Anything in images and videos. arXiv.
Siméoni, O., Vo, H. V., Seitzer, M., Baldassarre, F., Oquab, M., Jose, C., Khalidov, V., Szafraniec, M., Yi, S., Ramamonjisoa, M., Massa, F., Haziza, D., Wehrstedt, L., Wang, J., Darcet, T., Moutakanni, T., Sentana, L., Roberts, C., Vedaldi, A., Tolan, J., Brandt, J., Couprie, C., Mairal, J., Jégou, H., & Labatut, P. (2025). DINOv3. arXiv. https://doi.org/10.48550/arXiv.2508.10104
Xiong, Z., Wang, Y., Zhang, F., Stewart, A. J., Hanna, J., Borth, D., Papoutsis, I., Saux, B. L., Camps-Valls, G., & Zhu, X. X. (2024). Neural plasticity-inspired multimodal foundation model for earth observation. arXiv.