

AwakeForest: An Interactive Geospatial Platform for Large-Scale Forest Imagery

Suraj Prasai¹, Kangning Cui¹, Rongkun Zhu², Sarra Alqahtani¹, Ying Zhang¹,
Victor Paul Pauca¹, Miles R. Silman¹ and Fan Yang¹

¹Wake Forest University

²Hong Kong Baptist University

{prass25, cuij, alqahtas, zhangyi, paucavp, silmanmr, yangfan}@wfu.edu, csrkzhu@comp.hkbu.edu.hk

Abstract

Forest imagery analysis often involves multiple tightly coupled vision tasks, which must be performed under substantial variation in geographic regions, sensors, and acquisition conditions. However, practitioners often lack a *unified* tool that is geospatial-native, cloud-optimized, and ML-integrated for end-to-end workflows spanning annotation, prediction, visualization, and downstream analysis at scale. We present AwakeForest, an interactive end-to-end platform designed for large-scale forest imagery that integrates model-assisted inference, automatic annotation, and human-in-the-loop refinement within a single workflow. Our platform supports *plug-and-play* integration of pretrained models and enables scalable interaction with forest imagery ranging from standard aerial scenes to large orthomosaics that can span several gigabytes to hundreds of gigabytes. AwakeForest produces analysis-ready outputs that can be directly used for downstream analysis and to support iterative model and annotation updates on new scenes. We demonstrate the system on the PALMS dataset and illustrate how AwakeForest supports an end-to-end workflow for practical forest management and analysis.

1 Introduction

Forest imagery analysis commonly requires multiple vision tasks, including object detection, spatial delineation, and quantitative analysis, for ecological monitoring and forest management [Camalan *et al.*, 2022; de Almeida *et al.*, 2025; Cui *et al.*, 2025a]. In practice, these tasks are tightly coupled and need to be conducted within a single workflow. However, constructing high-quality labeled datasets is costly, and domain shifts across regions, sensors, and acquisition conditions further limit the reuse of existing annotations and models. As a result, practitioners often need to repeatedly adapt models and rebuild datasets with limited integrated tooling.

Several existing platforms support model-assisted labeling and integration of pretrained models. Tools such as CVAT [CVAT.ai Corp., 2023], AnyLabeling [Nguyen, 2025], and Label Studio [Tkachenko *et al.*, 2020] allow users to load

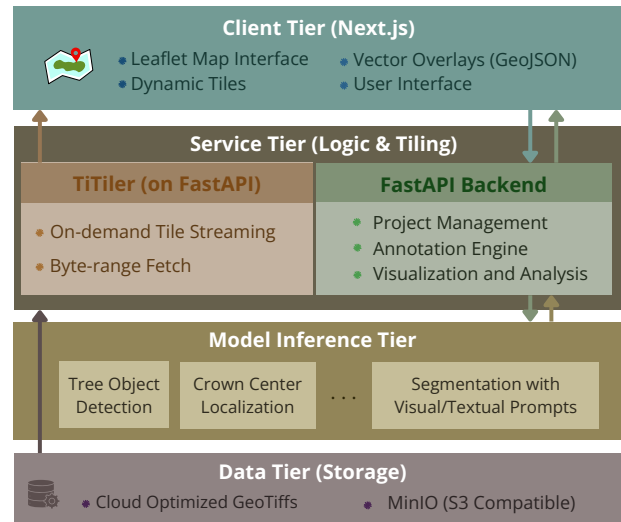


Figure 1: Overview of AwakeForest design.

detection or segmentation models to assist manual annotation and are effective for generic object annotation workflows and conventional computer vision datasets. However, they are not specifically designed for forest and ecological analysis, where imagery is often high-resolution and spatially extensive. In many scenarios, ecologists work with large orthomosaic images produced by automated reconstruction pipelines, where a single image can range from several to hundreds of gigabytes and objects of interest may be small and densely distributed. In these settings, annotation is tightly coupled with analysis, requiring iterative integration of detection, segmentation, counting, and domain-driven refinement rather than treating labeling as a standalone preprocessing step.

AwakeForest is built to address this gap by focusing particularly on large-scale forest imagery and practical forest management. We provide a unified interactive design (Figure 1) that integrates automated predictions with human-in-the-loop refinement, allowing results to be reused for downstream analysis or model adaptation across multiple tasks.

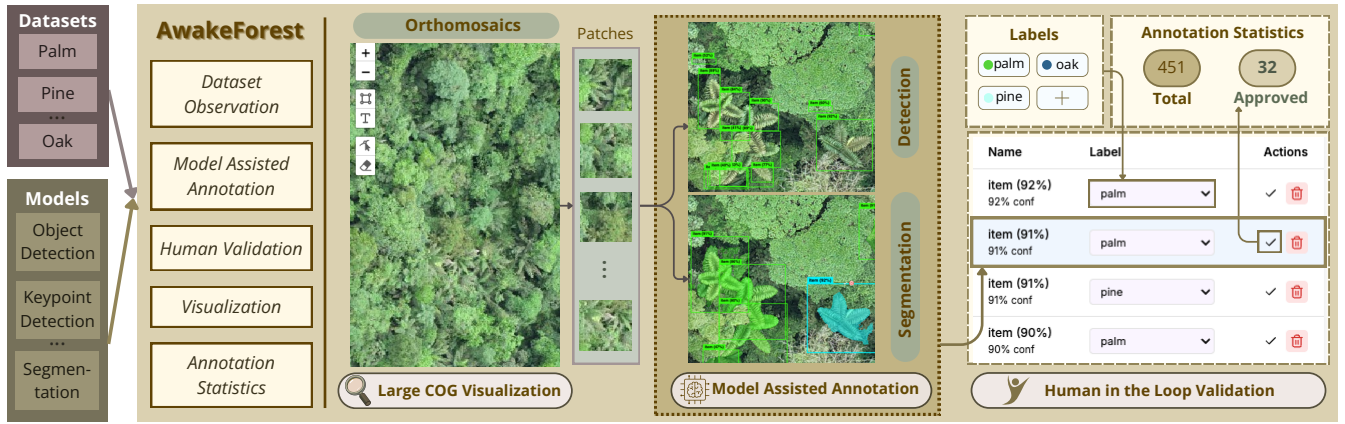


Figure 2: The AwakeForest workflow and interface. *Left*: Project configuration panel where users link dataset and select registered models. *Right*: Annotation workflow showing: large COG orthomosaic in global view, patch-based subdivision, model-assisted annotation using YOLO and YOLO+SAM2, human-in-the-loop label assignment and verification, and annotation statistics.

2 System Overview

2.1 High-Level System Architecture

AwakeForest adopts a modular, service-oriented architecture that separates user interaction, workflow coordination, model inference, and data management (see Figure 1). This separation allows system components to evolve independently while maintaining a coherent end-to-end workflow.

The *client tier* provides a user interface for managing datasets, models, and projects. Users can register datasets by specifying object storage endpoints and bucket details, add detection or segmentation models hosted locally or remotely by providing API endpoints, and create projects that link datasets with models. The *service tier* serves as a central coordinator responsible for user authentication, project organization, annotation management, and orchestration of inference requests across tasks. The *model inference tier* consists of independent services registered by users through model endpoints and task type specifications. The *data tier* serves large datasets through scalable object storage with on-demand tile streaming, separating raw imagery from spatial metadata and annotations to support interactive analysis.

2.2 Integrated Geospatial Annotation Workflow

AwakeForest organizes forest imagery analysis through a project-based structure that links geospatial datasets with external model services for interactive annotation, as shown in Figure 2. Each project connects a dataset with the registered models and provides two complementary views of the data.

The *global view* displays the entire orthomosaic, allowing users to navigate with pan and zoom controls while viewing annotation overlays for spatial context and progress tracking across large forest scenes in real time. The *patch-level annotation view* divides the dataset into fixed-size image patches where users can select models from a drop-down menu to generate predictions on the current patch; manually select and annotate objects missed by the models; assign labels and verify annotations for quality control and consistency; and navigate between patches to systematically annotate the dataset.

To ensure consistency across automated predictions and user edits, AwakeForest maintains a unified internal representation of intermediate results based on geospatial coordinates. By decoupling annotations from specific image tiles and patch boundaries, objects of interest are stored with absolute spatial attributes and standardized geometry formats. This design supports stable indexing, cross-model comparison, and incremental updates during refinement. It allows annotations to be managed and overlaid independently across diverse datasets or imagery sources with different resolutions, enabling outputs from various models to be combined, inspected, and refined within a cohesive spatial context.

2.3 Implementation Details

The platform is implemented using Next.js [Vercel Inc., 2026] for the *client tier* and FastAPI [Ramírez, 2024] for the *service tier*, with Leaflet [Leaflet Contributors, 2023] providing interactive geospatial map visualization. Data persistence and operations are managed through Supabase [Supabase Inc., 2026], which provides spatial indexing via PostGIS (a PostgreSQL extension for geospatial queries) [Refractons Research, 2024] and JWT-based authentication (JSON Web Tokens for secure, stateless session management).

Cloud Optimized GeoTIFFs (COGs) [Community, 2023] are stored in MinIO [MinIO Inc., 2025] object storage and served through dynamic tiling via TiTiler [DevelopmentSeed, 2025] integration. The service tier handles on-demand tile generation based on client viewport requests, enabling efficient streaming of image data at the scale required for interactive exploration of large datasets without preprocessing.

Model inference services are registered as external FastAPI endpoints defined by task type, such as detection or segmentation. This approach allows the platform to invoke remote inference during patch-level annotation without coupling inference logic to the core application. The architecture supports integration of various vision models including YOLO [Redmon et al., 2016; Jocher et al., 2023] detectors and SAM [Ravi et al., 2025; Carion et al., 2026] models for object detection, segmentation, and automatic labeling tasks.

Tool	Geospatial Native	Spatial Indexing	Cloud Optimized	AI-Assisted Labeling
Label Studio	×	×	×	✓
CVAT	×	×	×	✓
AnyLabeling	×	×	×	✓
QGIS	✓	✓	✓	×
AwakeForest	✓	✓	✓	✓

Table 1: Comparison with popular annotation tools.

2.4 Comparison with Existing Tools

Existing annotation tools such as Label Studio, AnyLabeling, and CVAT support generic image-centric labeling with optional AI-assisted proposals, but they do not treat geospatial coordinates and large orthomosaics as first-class primitives. Conversely, geographic information system (GIS) platforms such as QGIS [QGIS Development Team, 2026] are geospatial-native environments designed for spatial data management, indexing, and map-based analysis, yet they lack an integrated interactive ML annotation workflow with plug-and-play model inference capabilities. AwakeForest combines geospatial-first data management, cloud-optimized orthomosaic access, and model-integrated interactive labeling within a unified system architecture and workflow (Table 1).

3 Interactive Workflow and Demonstration

We show the capabilities of AwakeForest on the PALMS dataset [Cui *et al.*, 2025b], which provides a realistic case for interactive forest imagery analysis at scale. PALMS consists of high-resolution UAV-derived orthomosaics covering large, spatially continuous forest regions with dense and heterogeneous canopy structure. These characteristics require joint handling of labeling, object detection and localization, segmentation, and spatial reasoning, which are well suited for illustrating how AwakeForest supports end-to-end interaction from initial inference to refined, analysis-ready results.

Within AwakeForest, PALMS is loaded as large orthomosaics and interactively explored at multiple spatial scales. Users first identify regions of interest corresponding to forest patches or survey areas and initiate object searching using pretrained detection models. These models generate initial proposals for palm instances, including bounding boxes and candidate crown centers. Optionally, users can invoke SAM-based segmentation using these predicted boxes as prompts to obtain high-resolution instance masks for each detected palm. This enables quick conversion from coarse localization to detailed crown delineation without manual mask drawing.

Users can also invoke prompt-based segmentation for automatic labeling using either bounding boxes or text prompts. For PALMS, text prompts such as “palm” or “palm tree” can be used to retrieve candidate instances and generate corresponding masks for rapid initialization. With box prompts, the system produces a mask for the specified region and leverages the few-shot capability of SAM3 to suggest additional masks for visually similar crowns beyond the initial box. This interaction supports rapid discovery and correction of missed instances and boundary errors, reducing the need for manual annotation from scratch and improving overall efficiency.

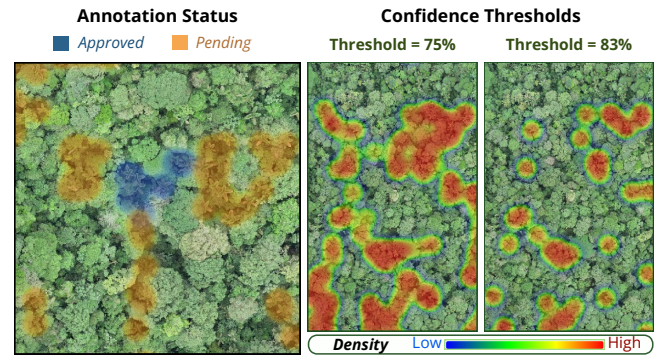


Figure 3: The analysis-ready outputs. Orthomosaic-level views for status verification and confidence filtering; prediction densities are computed via kernel density estimation over instance locations.

Throughout the demonstration, automated predictions and user refinements are maintained within a single project context. Verified results can be promoted to finalized annotations, while intermediate predictions remain traceable to their originating models. Figure 3 summarizes the analysis-ready outcomes at the orthomosaic level, combining verification status with confidence-thresholded prediction densities. Users can filter predictions by confidence, and the retained instance locations are summarized via kernel density estimation to reveal spatial clusters and coverage patterns. The finalized outputs are then used to derive spatial summaries such as palm center distributions and instance counts over selected regions [Cui *et al.*, 2025a; Zhu *et al.*, 2026].

4 Discussion and Conclusion

AwakeForest is an interactive platform that integrates model-assisted inference with human-in-the-loop refinement for large-scale forest imagery analysis. As a workflow-oriented platform, the quality of intermediate and final results depends on the suitability of the employed pretrained models, particularly under substantial domain shift or visually complex forest conditions, where additional user refinement may be required [Tuia *et al.*, 2016; Quiñero-Candela *et al.*, 2008]. Interactive analysis over very large orthomosaics also entails practical computational considerations [Fails and Olsen, 2003; Amershi *et al.*, 2014], such as inference latency and visualization overhead, which are inherent to large-scale geospatial data processing [Gorelick and others, 2017].

This work focuses on demonstrating the platform on RGB forest imagery in a representative forest management scenario and does not aim to provide quantitative benchmarks on predictive accuracy or annotation efficiency. Extensions to additional sensing modalities, including multispectral imagery [Hansen *et al.*, 2013], LiDAR [Lefsky *et al.*, 2002], and temporal data [Banskota *et al.*, 2014], as well as collaborative and multi-user workflows [Palomino *et al.*, 2017; Ma *et al.*, 2022], are left as future work. Within its current scope, AwakeForest illustrates how an integrated, end-to-end workflow can support detection, segmentation, spatial summarization, and iterative refinement in a unified environment for practical forest monitoring and analysis.

Acknowledgments

We thank the reviewers for their constructive feedback. We also thank the Center for Amazonian Scientific Innovation (CINCIA) for providing access to the data used in this work. The project source code is available at <https://github.com/spygaurad/AwakeForest>, and the demonstration video can be viewed at <https://youtu.be/Z0EZj1PPDv4>.

Contribution Statement

Suraj Prasai and Kangning Cui contributed equally to this work. Fan Yang served as the corresponding author.

References

- [Amershi *et al.*, 2014] Saleema Amershi, Maya Cakmak, W. Bradley Knox, and Todd Kulesza. Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4):105–120, 2014.
- [Banskota *et al.*, 2014] Asim Banskota, Nilam Kayastha, Michael J. Falkowski, et al. Forest monitoring using land-sat time series data: A review. *Canadian Journal of Remote Sensing*, 40(5):362–384, 2014.
- [Camalan *et al.*, 2022] Seda Camalan, Kangning Cui, Victor Paul Pauca, et al. Change detection of amazonian alluvial gold mining using deep learning and sentinel-2 imagery. *Remote Sensing*, 14(7):1746, 2022.
- [Carion *et al.*, 2026] Nicolas Carion, Laura Gustafson, et al. SAM 3: Segment anything with concepts. In *International Conference on Learning Representations*, 2026.
- [Community, 2023] Cloud Optimized GeoTIFF Community. Cloud Optimized GeoTIFF. <https://www.ogc.org/standards/geotiff>, 2023.
- [Cui *et al.*, 2025a] Kangning Cui, Wei Tang, Rongkun Zhu, et al. Efficient localization and spatial distribution modeling of canopy palms using uav imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
- [Cui *et al.*, 2025b] Kangning Cui, Rongkun Zhu, et al. Detection and geographic localization of natural objects in the wild: A case study on palms. In *International Joint Conference on Artificial Intelligence*, 2025.
- [CVAT.ai Corp., 2023] CVAT.ai Corp. Computer Vision Annotation Tool. <https://github.com/cvat-ai/cvat>, 2023.
- [de Almeida *et al.*, 2025] Danilo RA de Almeida, Laura B Vedovato, et al. Remote sensing approaches to monitor tropical forest restoration: Current methods and future possibilities. *Journal of Applied Ecology*, 2025.
- [DevelopmentSeed, 2025] DevelopmentSeed. TiTiler. <https://github.com/developmentseed/titiler>, 2025.
- [Fails and Olsen, 2003] Jerry Alan Fails and Dan R. Olsen. Interactive machine learning. In *Proceedings of the 8th International Conference on Intelligent User Interfaces (IUI '03)*, pages 39–45. ACM, 2003.
- [Gorelick and others, 2017] Noel Gorelick et al. Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27, 2017.
- [Hansen *et al.*, 2013] M. C. Hansen, P. V. Potapov, et al. High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160):850–853, 2013.
- [Jocher *et al.*, 2023] Glenn Jocher, Jing Qiu, and Ayush Chaurasia. Ultralytics YOLO, January 2023.
- [Leaflet Contributors, 2023] Leaflet Contributors. Leaflet. <https://github.com/Leaflet/Leaflet>, 2023.
- [Lefsky *et al.*, 2002] Michael A. Lefsky, Warren B. Cohen, David J. Harding, Geoffrey G. Parker, Steven A. Acker, and S. Thomas Gower. Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography*, 11(5):393–399, 2002.
- [Ma *et al.*, 2022] Zaiyang Ma, Min Chen, Zhong Zheng, et al. Customizable process design for collaborative geographic analysis. *GIScience & Remote Sensing*, 59(1):914–935, 2022.
- [MinIO Inc., 2025] MinIO Inc. MinIO Object Storage. <https://github.com/minio/minio>, 2025.
- [Nguyen, 2025] Viet Anh Nguyen. AnyLabeling - Effortless data labeling with AI support. <https://github.com/vietanhdev/anylabeling>, 2025.
- [Palomino *et al.*, 2017] Jenny Palomino, Oliver C Muellerklein, and Maggi Kelly. A review of the emergent ecosystem of collaborative geospatial tools for addressing environmental challenges. *Computers, Environment and Urban Systems*, 65:79–92, 2017.
- [QGIS Development Team, 2026] QGIS Development Team. Qgis geographic information system. <https://www.qgis.org>, 2026.
- [Quiñonero-Candela *et al.*, 2008] Joaquin Quiñonero-Candela, Masashi Sugiyama, Anton Schwaighofer, and Neil D. Lawrence, editors. *Dataset Shift in Machine Learning*. The MIT Press, Cambridge, MA, 2008.
- [Ramírez, 2024] Sebastián Ramírez. FastAPI. <https://fastapi.tiangolo.com>, 2024.
- [Ravi *et al.*, 2025] Nikhila Ravi, Valentin Gabeur, et al. Sam 2: Segment anything in images and videos. In *International Conference on Learning Representations*, 2025.
- [Redmon *et al.*, 2016] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 779–788, 2016.
- [Refractions Research, 2024] Refractions Research. PostGIS. <https://postgis.net>, 2024.
- [Supabase Inc., 2026] Supabase Inc. Supabase. <https://supabase.com>, 2026.
- [Tkachenko *et al.*, 2020] Maxim Tkachenko, Mikhail Malyuk, Andrey Holmanyuk, and Nikolai Liubimov. Label Studio: Data labeling software. <https://github.com/HumanSignal/label-studio>, 2020.
- [Tuia *et al.*, 2016] Devis Tuia, Claudio Persello, and Lorenzo Bruzzone. Domain adaptation for the classification of remote sensing data: An overview of

recent advances. *IEEE Geoscience and Remote Sensing Magazine*, 4(2):41–57, 2016.

[Vercel Inc., 2026] Vercel Inc. Next.js 14. <https://nextjs.org>, 2026.

[Zhu *et al.*, 2026] Rongkun Zhu, Kangning Cui, et al. From orthomosaics to raw uav imagery: Enhancing palm detection and crown-center localization. *International Geoscience and Remote Sensing Symposium*, 2026.