

Reducing expert annotation effort for drone-based plant detection with vision foundation models

Starting date: as soon as possible

Duration: 6-12 months

Working Place: Agroscope Posieux

Language: French/English

Contact: Hassan-Roland Nasser, Helge Aasen

Email: roland.nasser@agroscope.admin.ch, helge.aasen@agroscope.admin.ch

Background

Training a detection model for a new plant species or a new field site still depends on hours of manual annotation by experts. Vision foundation models such as SAM 3 (Meta, 2025) and Grounding DINO can now produce a first set of labels from a text prompt or a few example images, and tools like Autodistill or ARAMSAM (Oehme et al., 2026) already use them to speed up labelling. The open question is where these automatic labels can be trusted and where an expert is still needed. Published results show that foundation models work well for large, common objects but struggle with visually similar species and small objects, and that detectors trained on one site or one month of the flowering season can lose most of their accuracy at another (Gallmann et al., 2022; Frontiers in Plant Science, 2025).

Project

The thesis is embedded in an Agroscope project on flowering field margins, in which 25 margins in three cantons are flown by drone roughly monthly from April to August to estimate the flower resources available to pollinators under different mowing regimes. A second Master student with botanical training, based at Reckenholz, will provide expert labels on a subset of the images; this thesis provides the computer vision side. Together the two theses build an annotated dataset of drone images of Swiss flowering plants.

The research questions, to be refined with the student, are:

- How much expert annotation does an iterative, semi-automatic labelling loop need to reach expert-level labels on a multi-species flower dataset, and how does this vary with species and flower size?
- Can the quality of automatically generated labels be estimated without ground truth (for example through agreement between models or prompts), and does that estimate predict real accuracy per species?
- How much does detection accuracy drop between months and between sites, and how much of it can be recovered by re-prompting with a few current examples?

The work covers the whole chain: drone image preparation, running and adapting foundation models, designing the labelling loop, evaluation against the botanist's labels, and writing up. The same methods will be tested on other public datasets to check that the findings transfer.

Since this is an active research field, the exact scope and methods may be adjusted during the project as the literature evolves.

What we offer

- Real field data with a clear ecological question behind it



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Economic Affairs,
Education and Research EAER
Agroscope

Swiss Confederation

- Close supervision and regular exchange with both the vision and the ecology teams
- Access to GPU computing, annotation tools and model tracking infrastructure
- A dataset and results intended for publication

Technical stack

Python, Git, Jupyter / VS Code, PyTorch, Annotation Tools, CSCS Cluster stack (SLURM, PyFireCrest, ...), MLFlow, NVIDIA GPUs, Linux

Prior knowledge

Good proficiency in Python and basic knowledge of deep learning. Experience with computer vision, drones or GIS is a plus; the rest can be learned during the project. Interest in plants is welcome but botanical expertise is not required.

References

- Gallmann et al., 2022. Flower mapping in grasslands with drones and deep learning. *Frontiers in Plant Science* 12:774965.
- Oehme et al., 2026. Orchestrating segment anything models to accelerate segmentation annotation on agricultural image datasets. *Frontiers in Artificial Intelligence* 8:1748468.
- Wang et al., 2026. Accurate weed detection in UAV grassland images via intelligent annotation. *Frontiers in Plant Science* 17.
- Meta AI, 2025. SAM 3: Segment Anything with Concepts.