Tropentag, September 16-18, 2026, hybrid conference

"Towards multi-functional agro-ecosystems
promoting climate resilient futures"

Foundation-model segmentation for automated annotation of plant species

Shubham Rana[1], Oliver Hensel[2], Abozar Nasirahmadi[3]

[1] *Universitat Kassel, Agricultural and Biosystems Engineering, Germany*
[2] *University of Kassel, Agricultural and Biosystems Engineering, Germany*
[3] *University of Kassel, Agricultural and Biosystems Engineering,*

Abstract

Manual annotation of vegetation in field-based rangeland imagery is slow, subjective, and difficult to standardise. These challenges are particularly acute in remote environments where complex plant structures and visually cluttered scenes limit the efficiency of traditional mask generation for training purposes, particularly due to spectral overlap and spectral heterogeneity. This research addresses these bottlenecks by introducing a generative semantic reconstruction method designed for the automated annotation of indigenous forage species. The study utilises proximally acquired RGB imagery sourced from diverse rangeland ecosystems in Marsabit County, Kenya, and various regions across Namibia.

The workflow reconstructs raw field images into annotation-ready priors using the Gemini Nano Banana Pro 2 model. To ensure class completeness for images with occlusions or poorly captured features, multiple perspectives of these priors were also generated to recover missing fine-scale details. These outputs then undergo automated mask generation with SAM 3, followed by validation using YOLOv8, YOLOv11, and RF-DETR. An ablation study was conducted to verify the reconstruction stage's robustness. Qwen Edit 2512 and ChatGPT 5.4 were used, applying identical prompting logic and evaluation protocols.

The analysis showed that the Gemini Nano Banana Pro 2 produced the most reliable results, demonstrating subject preservation and usable segmentation priors. This model also showed high downstream detection accuracy. This approach shifts the annotation burden from manual pixel-wise delineation to generative preconditioning coupled with foundation-model segmentation. This establishes a scalable framework for vegetation datasets in challenging environments. The research provides a transferable strategy for accelerating ecological computer vision workflows where conventional labeling is expensive.

Keywords: Generative AI , image-to-image translation. style transfer, pixel purity

Contact Address: Shubham Rana, Universitat Kassel, Agricultural and Biosystems Engineering, Fachgebietsleiter agrartechnik nordbahnhoftsrasse 1a zimmer 1140 max eyth haus, 37213 Witzenhausen, Germany, e-mail: shubham.rana@uni-kassel.de