



Foundation-model Segmentation for Automated Annotation of Plant Species

Shubham Rana ¹, Oliver Hensel ¹, Abozar Nasirahmadi ^{1,2}

¹ Department of Agricultural and Biosystems Engineering, University of Kassel, Witzenhausen, D-37213, Germany

² Department of Energy and Technology, Swedish University of Agricultural Sciences, Box 7032, Uppsala 75007, Sweden

Background & research questions

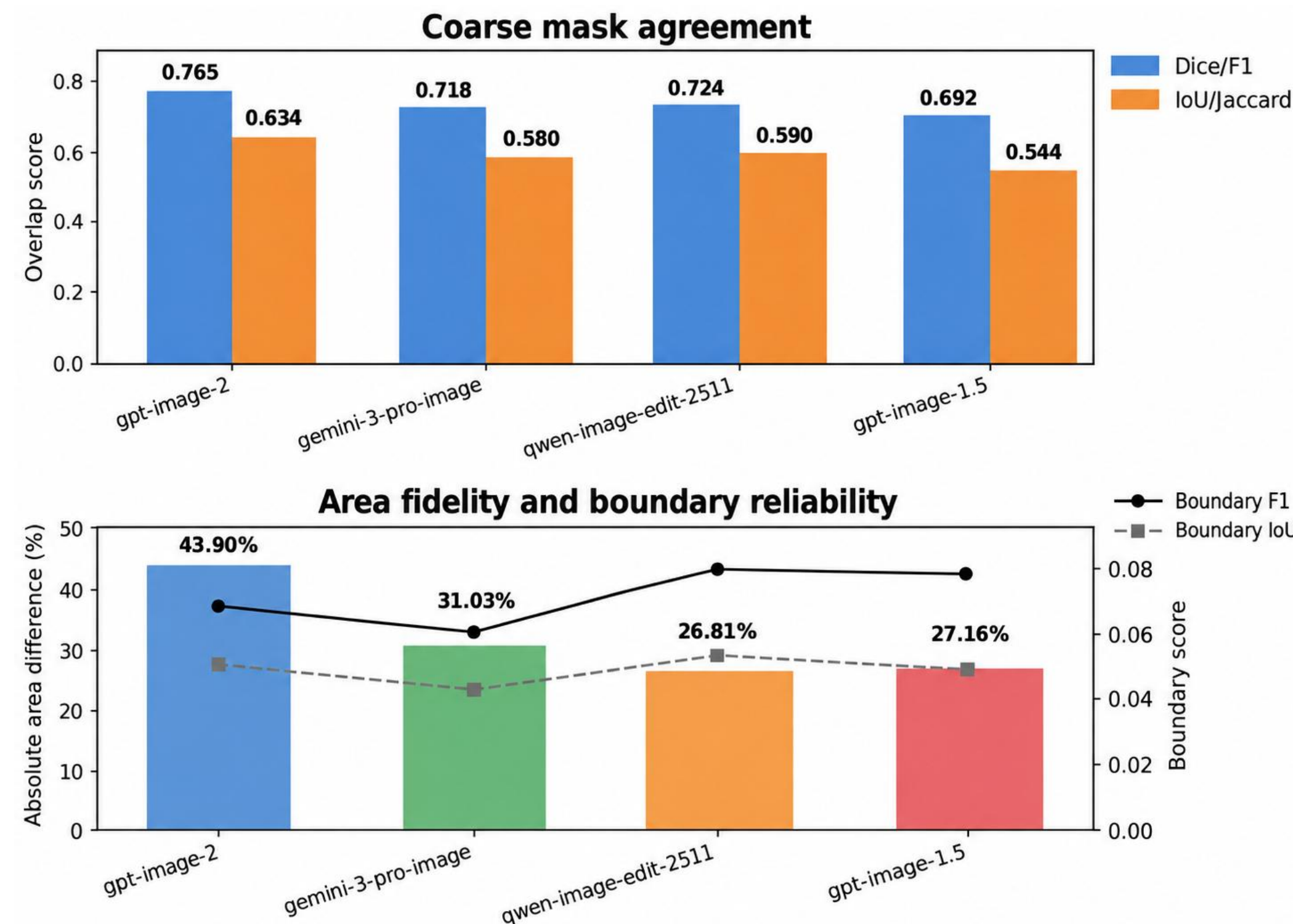
- **Annotation bottleneck.** Semi-arid RGB scenes combine sparse shrubs, exposed woody structures, senescent material, soil, litter, stones and shadows, producing strong vegetation to background ambiguity.
- **Generative opportunity and risk.** Foundation models can suppress clutter and accelerate annotation, but may delete fine branches, fill natural canopy gaps or add plant-like structures.
- **Core question.** Can a fixed global prompt produce annotation-ready vegetation priors while preserving biologically meaningful shrub morphology?

Research questions

- 1) Which model best preserves overlap, boundary, fine structure, colour and canopy gaps?
- 2) Do models fail through complementary mechanisms?
- 3) Can model disagreement be converted into a human-review signal?

Results

1 | Mask agreement & boundary reliability



• GPT-image-2 maximized overlap but also produced the largest absolute area difference (43.90%).

• Qwen-image-edit-2511 minimized area difference (26.81%) and achieved the highest Boundary F1 (0.075) / Boundary IoU (0.055).

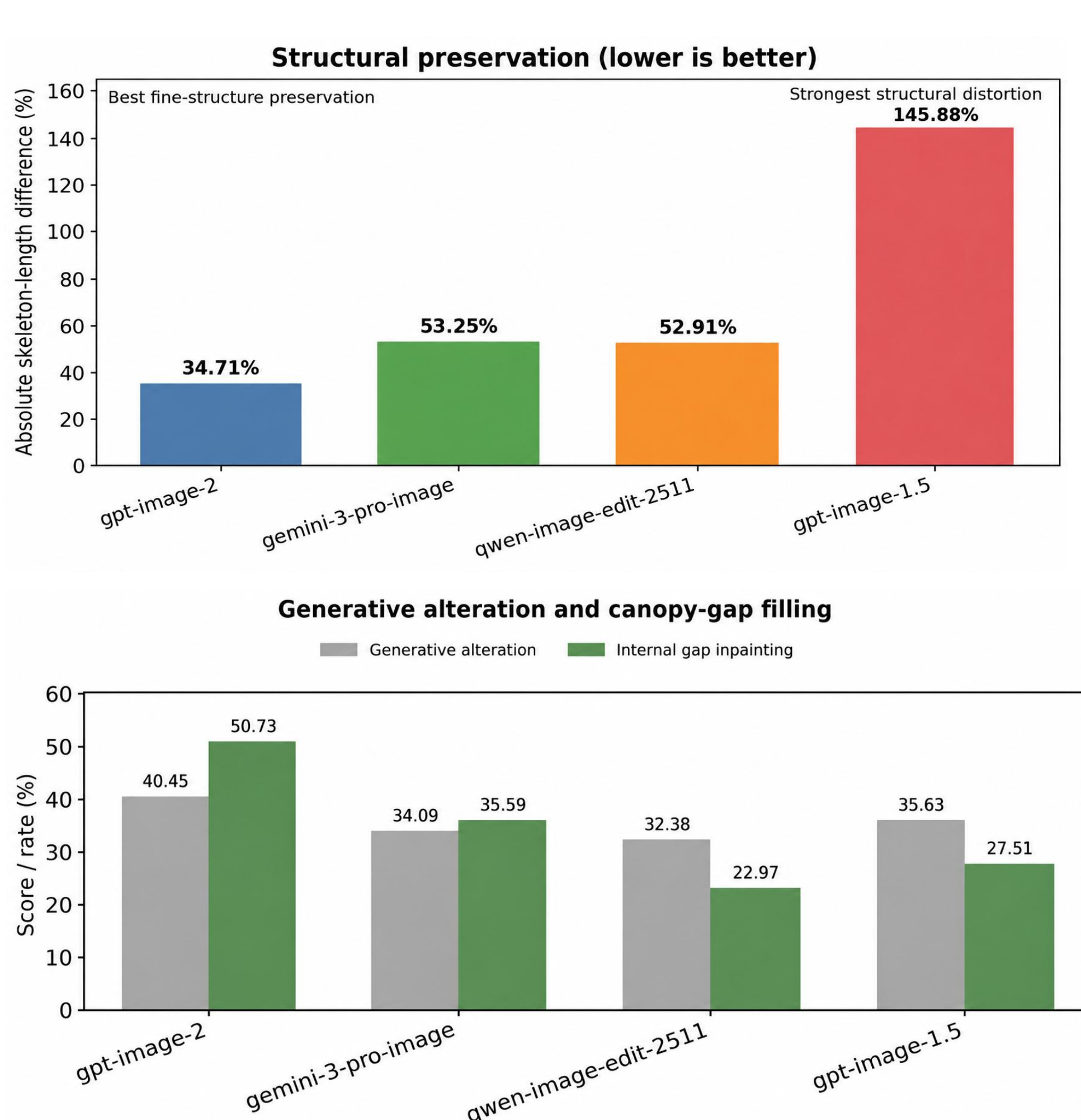
• Fine twigs, canopy gaps and vegetation–soil contacts remained difficult for every system.

Reliability dimension	Best model	Value
Overlap	GPT-image-2	Dice 0.765
Boundary	Qwen-image-edit-2511	F1 0.075
Colour	Gemini-3-pro-image	ΔE 25.61
SSIM	GPT-image-1.5	0.227

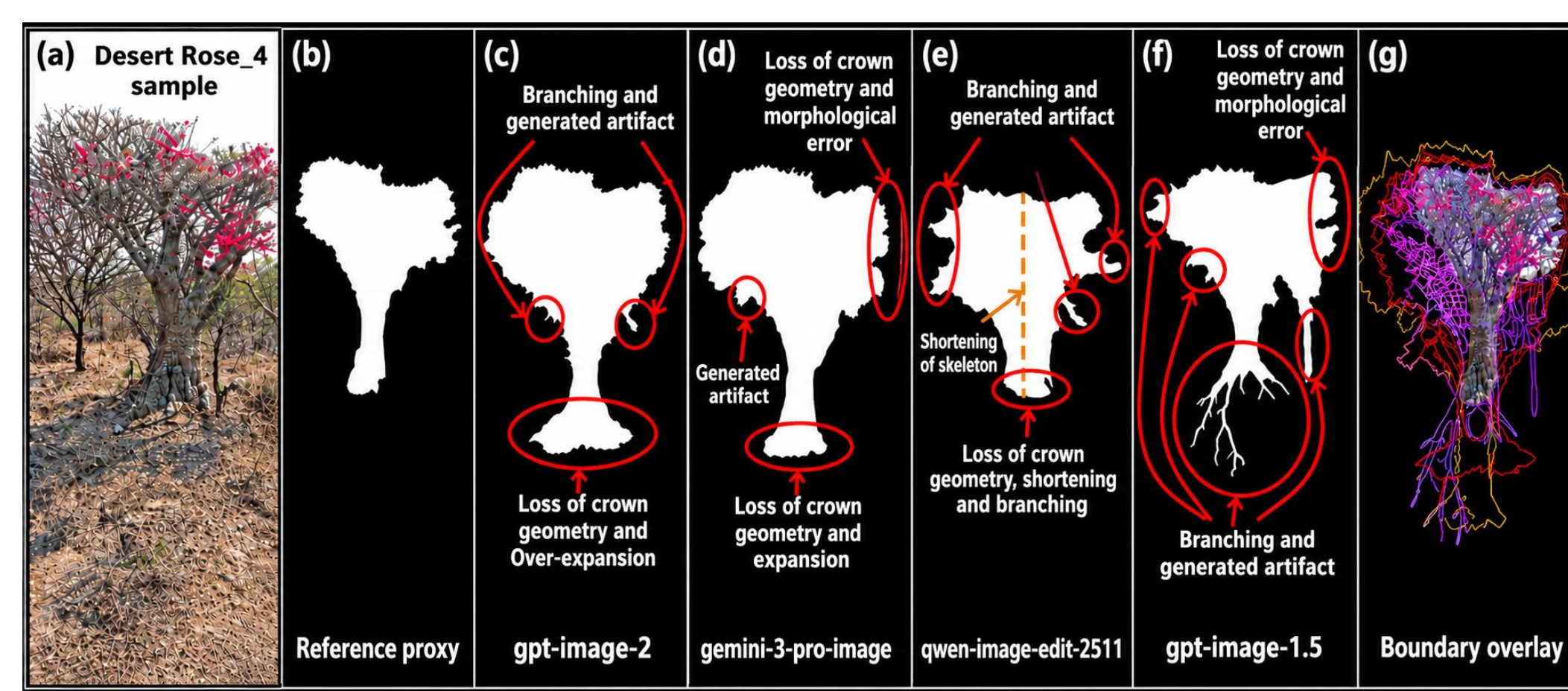
Interpretation:

Coarse overlap alone is insufficient to quantify segmentation-mask fidelity. Boundary- and architecture-sensitive metrics are essential for faithful annotations.

2 | Structural & generative failure modes



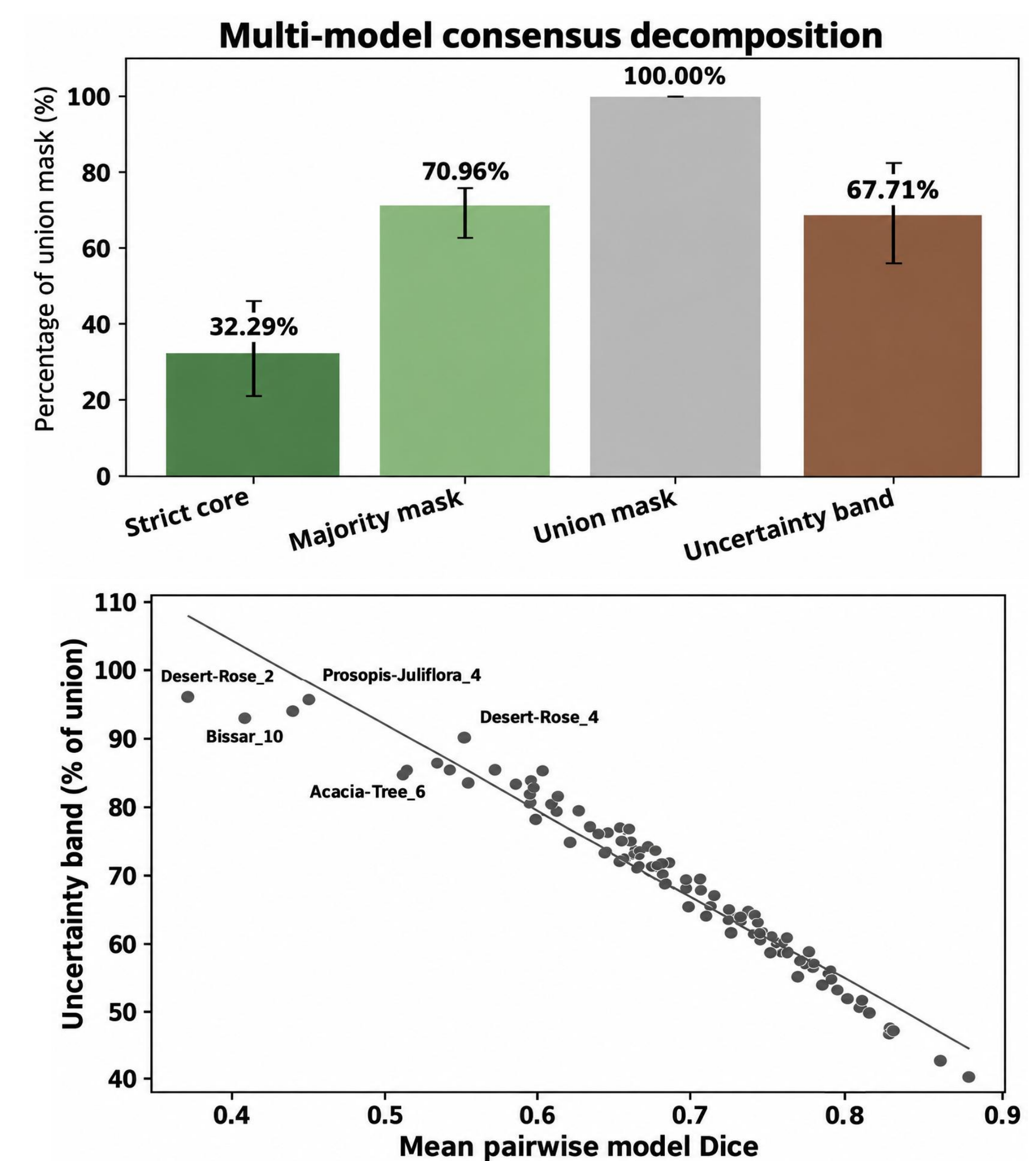
Representative failure modes



Desert Rose example: crown-geometry loss and expansion, branch deletion, skeleton shortening and generated artefacts. These errors are visually small but structurally important.

Model-specific trade-offs: GPT-image-2 preserved the skeleton best but filled most gaps; Qwen was more conservative; GPT-image-1.5 showed the strongest skeleton distortion (145.88%).

3 | Consensus uncertainty → human review



Highest priority human review samples

Sample	Class	Uncertainty	Pairwise Dice
Desert-Rose_2	Desert Rose	95.60%	0.371
Prosopis-Juliflora_4	Prosopis juliflora	95.21%	0.432
Bissar_10	Bissar	93.60%	0.407
Desert-Rose_4	Desert Rose	88.91%	0.536
Acacia-Tree_6	Acacia Tree	88.26%	0.513

Consensus decision rule

Strict core (all 4 models) → candidate high-confidence vegetation
Majority mask → moderate confidence
Uncertainty band → expert review / manual refinement
Mean pairwise Dice = 0.696 ± 0.094 ; IoU = 0.548 ± 0.104

Study area & dataset



Boscia coriacea

Bissar

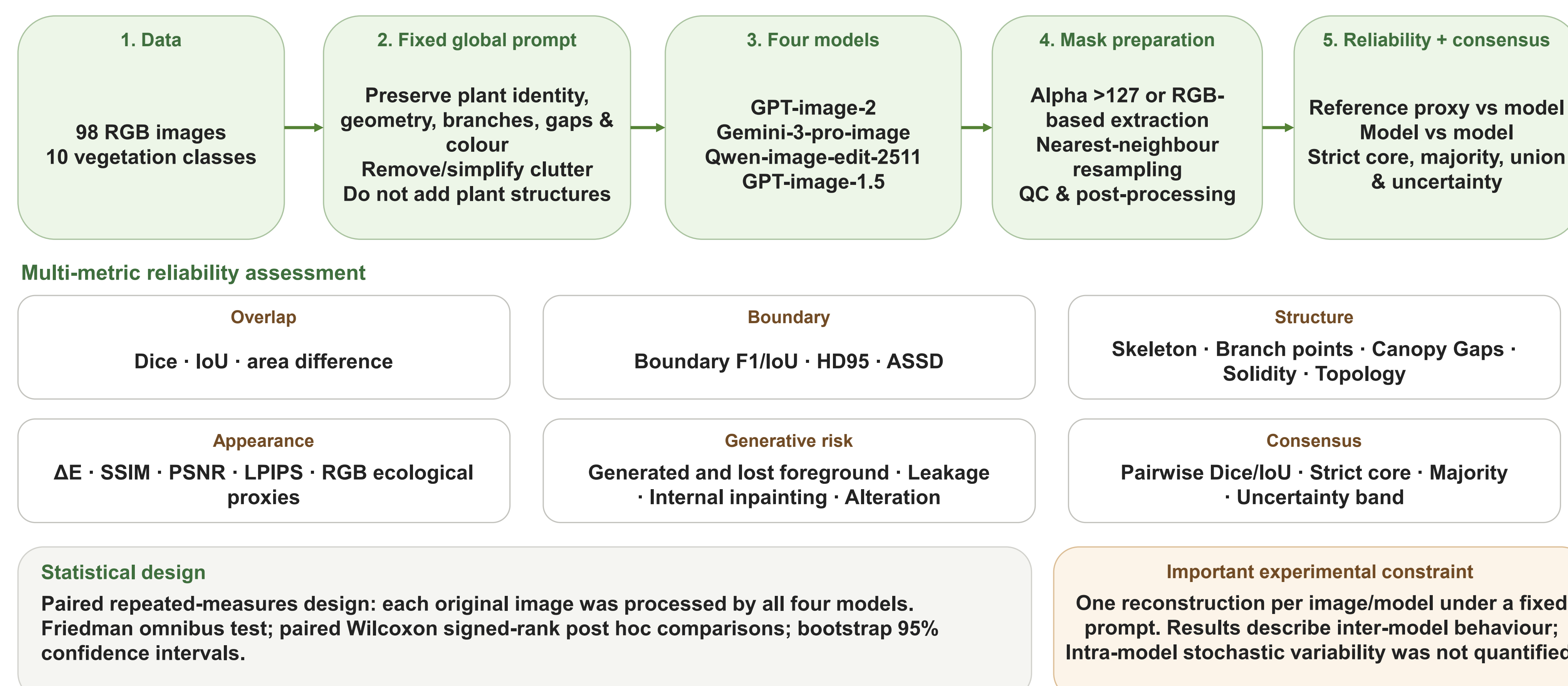
Desert Rose

- Location: Marsabit County, Kenya; proximal RGB imagery from rainy and dry seasons.
- 98 original images; 10 forage shrub/rangeland vegetation classes.
- 4 reconstructions per image → 392 reconstructed image–mask pairs.
- Common spatial grid: 1184 × 864 pixels.
- Visual challenges: open crowns, thorns, sparse woody branches, dry leaves, shadows, soil, stones, litter and background vegetation.

Classes: Acacia, Acacia Tree, Bissar, Boscia coriacea, Candelabra, Desert Rose, Duma, Maerua edulis, Prosopis juliflora and Yabah

Practical outcome: A reliability-screened human-in-the-loop annotation workflow for targeted annotation and segmentation in highly heterogeneous environments.

Methodology & statistical design



Selected references

Rana et al. (2026), MethodsX 16, 103967
Hosseini et al. (2026), Remote Sens. Appl.: Soc. Environ. 41, 101957
Cheng et al. (2021), CVPR, Boundary IoU
Wu et al. (2025), Qwen-image technical report.



Contact:
Dr. Shubham Rana
shubham.rana@uni-kassel.de
ORCID: 0000-0003-4691-699X
WhatsApp: +49-15511571245
Office: +49 5542981680

UNIKASSEL
VERSITÄT
ORGANIC
AGRICULTURAL
SCIENCES



Funded by:



Federal Ministry
of Research, Technology
and Space



Tropentag 2026

Need solutions, expertise and support in computer vision and high-throughput crop phenotyping?
You are at the right place.