UNDERGRADUATE THESIS · RESEARCH RESULTS

Grading *ojo de gallo* on a phone.

An on-device system that quantifies ***Mycena citricolor*** damage on coffee leaves, matching lab-grade models **20× its size**, fully offline.

Carly Chery · EARTH University, Guácimo, Costa Rica · Thesis **PG 26038** · July 2026

100%	**0.94**	**17 MB**	**12**
Grade accuracy, held-out (QWK 1.00)	Peak lesion **Dice**, segmentation	Deployed model, runs **100% offline**	**Models** benchmarked, 5 architectures

THE TASK

From a leaf photo to a severity grade

Ojo de gallo ("rooster's eye"), caused by the fungus *Mycena citricolor*, is a major cause of leaf loss in Central-American coffee. The system measures the fraction of leaf area lost to lesions and assigns one of four standard severity grades, the number agronomists act on.

G0 Healthy · ∼0% **G1** Mild · < 5% **G2** Moderate · 5–20% **G3** Severe · ≥ 20%

HOW IT WORKS

A five-step vision pipeline

01 Leaf photo → **02** Leaf isolation (U^2-Net) → **03** Lesion segmentation (DeepLabV3+) → **04** Severity % (lesion ÷ leaf) → **05** Grade G0–G3

The model traces the **necrotic lesion boundary** pixel by pixel and measures its share of the isolated leaf. That fraction, not a black-box label, is what produces the grade, so every result is **visually auditable**, an essential property for research-grade phenotyping.

MODEL BENCHMARK

The lightweight model holds its own

Nine segmentation models plus a transformer classifier, all on the same held-out test set. Ranked by lesion Dice: the **deployed 17 MB model** sits third, behind only a 106 MB and a 324 MB network.

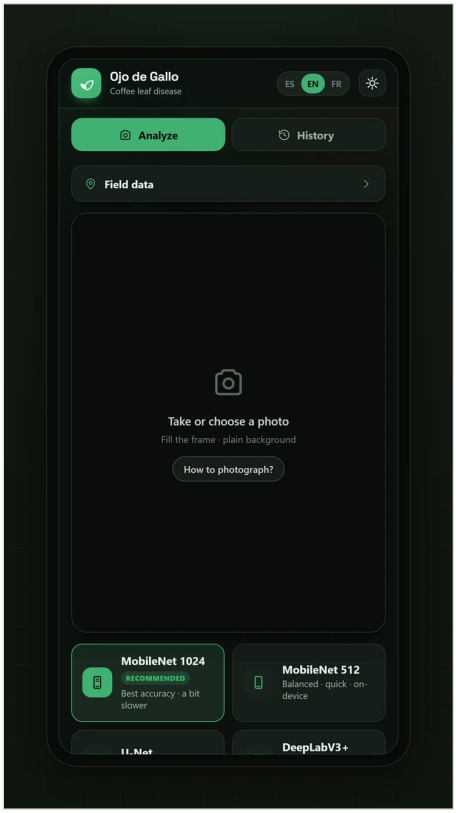
MODEL	GRADE	QWK	DICE	IoU	SIZE	ON-DEVICE
SegFormer-B5 (MiT-B5, 1024px)	100%	1.00	0.940	0.916	324 MB	–
DeepLabV3+ (ResNet-50, 1024px)	100%	1.00	0.933	0.906	106 MB	–
DeepLabV3+ MobileNetV2 (1024px)	**100%**	**1.00**	**0.924**	**0.895**	**17 MB**	**Yes**
SegFormer-B5 (MiT-B5, 512px)	100%	1.00	0.911	0.878	324 MB	–
DeepLabV3+ EfficientNet-B7 (512px)	96.4%	0.944	0.893	0.851	242 MB	–
DeepLabV3+ MobileNetV2 (512px)	98.2%	0.972	0.892	0.850	17 MB	**Yes**
DeepLabV3+ (ResNet-50, 512px)	98.2%	0.972	0.891	0.856	∼100 MB	–
PLA-ViT (ViT-B/16 classifier)	98.2%	0.972	n/a	n/a	328 MB	–

Per-leaf metrics on the real held-out test set (worst-face aggregation). Grade accuracy saturates at 100% for several models, so **lesion Dice and model size** are the true differentiators. PLA-ViT is a whole-leaf classifier and produces no lesion map.

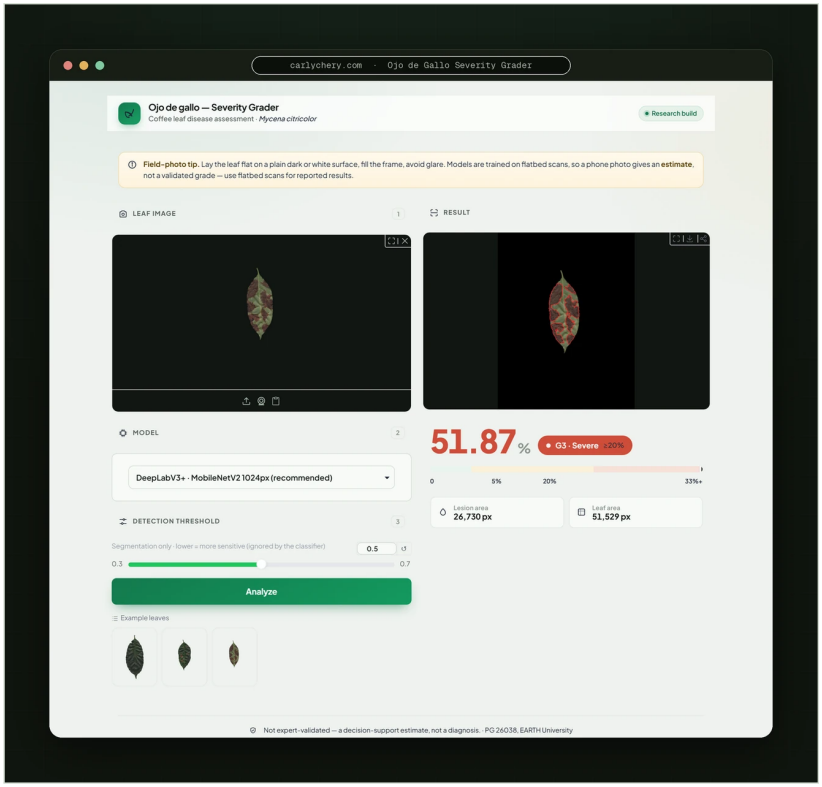
TWO WORKING PRODUCTS

A field app and a web tool, from the same models

Every model is exported to ONNX and runs in two shipped interfaces: an offline Android app for the field, and a desktop web tool for the lab.



On-device Android app: no connectivity, GPS geotagging, survey sessions, three languages, CSV/PDF export.



Desktop web tool: a severe leaf read at **51.9%** · **G3** by the deployed MobileNet model, with lesion overlay and pixel areas.

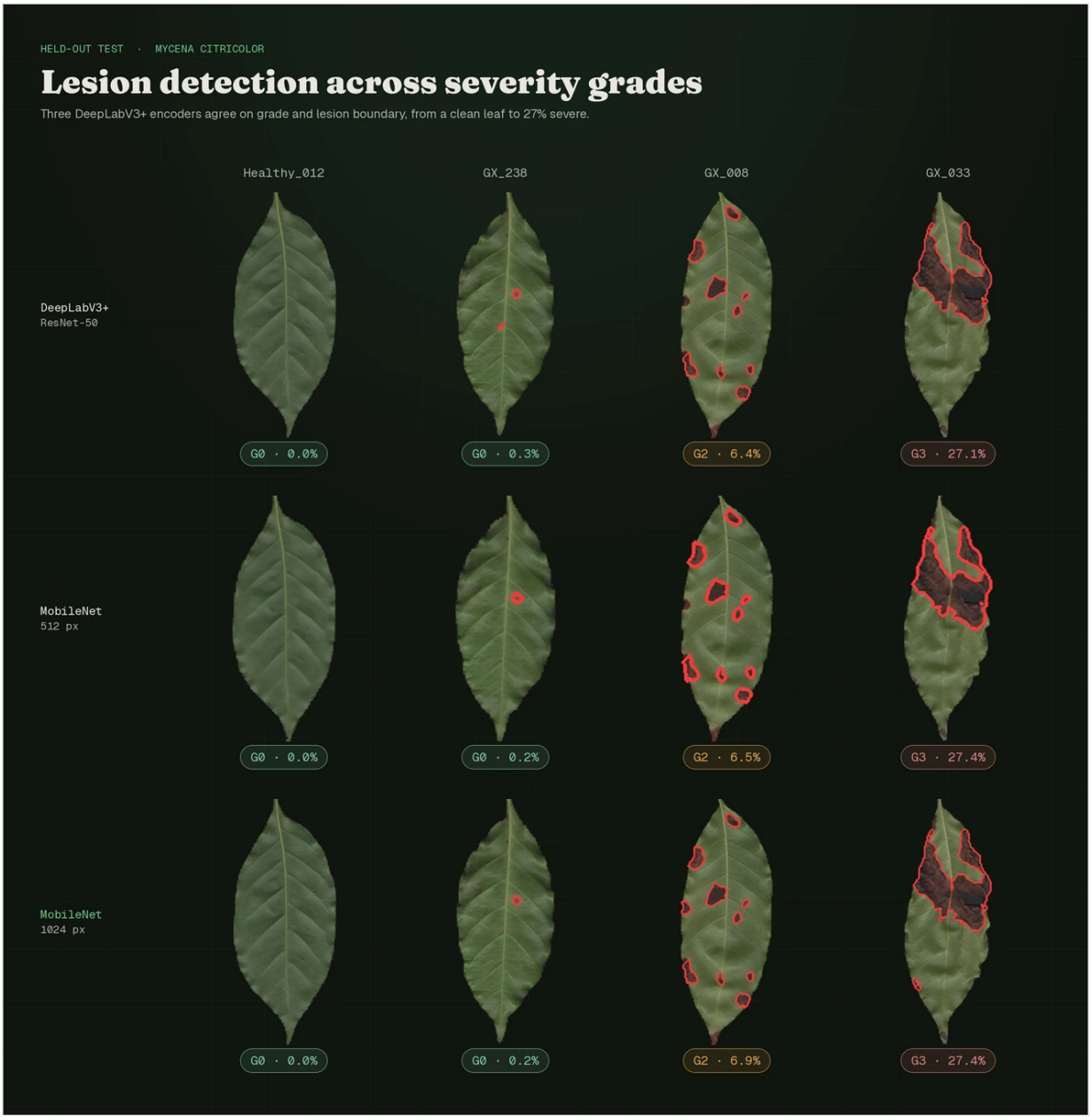
WHAT THE RESULTS SHOW

Three findings

1. **Small beats large.** A 17 MB MobileNet matched or beat ResNet-50 (106 MB), EfficientNet-B7 (242 MB), SegFormer-B5 (324 MB) and a ViT classifier (328 MB) on grade, and ranks third on lesion Dice. On this dataset, capacity did not help; efficiency did.

2. **Two paradigms converge.** Direct ViT classification and segment-then-measure reached identical grade accuracy (98.2%, QWK 0.972). But only segmentation yields an interpretable severity map and a defensible percentage, what a researcher can actually verify.

3. **Field-ready, not lab-bound.** The whole system runs offline on a phone (ONNX-Runtime WASM, no server, no signal) with GPS geotagging, survey-session incidence tracking, three languages and CSV/PDF export built for real field surveys.



Four leaves span G0 to G3; rows are DeepLabV3+ with ResNet-50, MobileNet-512 and MobileNet-1024 encoders. The lightweight MobileNet, the heavy ResNet-50 and the giant EfficientNet-B7 detect the same lesions and assign the same grades, from a clean healthy leaf to a 27%-severe one.

DATASET & METHODS

Balanced dataset of expert-graded coffee-leaf scans: **749 training** (real + synthetic augmentation), **110 validation** and **110 held-out test** images across three severity classes. Architectures: DeepLabV3+ (MobileNetV2 / ResNet-50 / EfficientNet-B7 encoders), SegFormer-B5, and a ViT-B/16 classifier, trained in PyTorch, exported to ONNX, and run on-device with onnxruntime-web (WASM). Leaf isolation uses a U^2-Net salient-object model; severity = lesion pixels ÷ leaf pixels.

A decision-support research tool for coffee-leaf disease assessment, not a validated clinical diagnosis. Reported metrics are on flatbed scans; phone photos yield estimates. Model artifacts, training code and full results available on request.