Efficient Deep Learning and Spatial Modeling for Mapping Tropical Forest Canopies

Jason Cui

Department of Computer Science
Wake Forest University

November 3, 2025

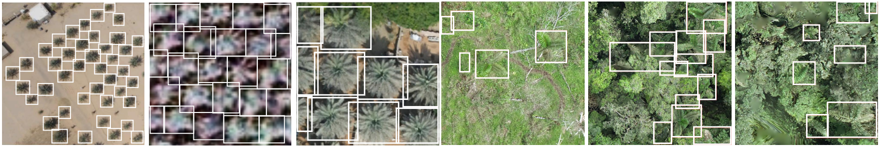
Self-Introduction

- **Positions**:
  Research Assistant Professor, Department of Computer Science;
  Robert J. Plemmons Fellow, Sabin Center for Environment & Sustainability.
- **Research**:
  Efficient machine learning; statistical & spatial modeling; interpretable deep learning (applications: remote sensing, ecology, healthcare).
- **Education**:
  Ph.D. in Mathematics, City University of Hong Kong.
  B.Sc. in Computing Mathematics, City University of Hong Kong.

Outline

Motivation and Problem

The Importance of Palms in Tropical Forests



(a) Cases from existing studies (b) Cases from our dataset

Figure: Comparative Samples of Manual Labels.

Ecological & Economic Significance of Palms:

- Vital to tropical forest ecology, biodiversity, and conservation planning [1].
- Support sustainable livelihoods and are key resources for tropical wildlife [2].
- Can serve as bioindicators of forest health and environmental impact.

Problem: Most prior work targets plantations; *natural forests* show irregular spacing, overlapping crowns, complex backgrounds, uneven lighting.

Research Challenges & Our Contributions

Key Challenges

- **Image Variability:** Occlusion from overlapping canopies and inconsistent lighting degrade image quality.
- **Data Scarcity:** High-quality annotated datasets for tropical forests are rare and difficult to create.
- **Spatial Analysis Gap:** Most work focuses only on detection, not the large-scale spatial structure of populations.

Our Contributions

1. **New Dataset (PALMS):** Including 21 forest sites' orthomosaics in western Ecuador, with over 8,800 bounding boxes and 5,000 palm center annotations.
2. **PRISM Framework:** A flexible, interpretable pipeline for palm detection, segmentation, and counting.
3. **Spatial Modeling:** A Poisson-Gaussian model that simulates and provides insight into the ecological processes driving palm distribution.

The PALMS Dataset

The PALMS Dataset: Data from the Field

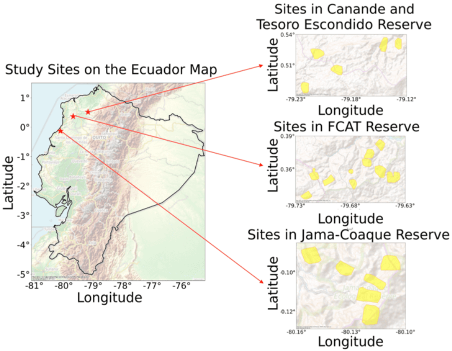


Figure: Study sites across a rainfall and ecological gradient in western Ecuador.

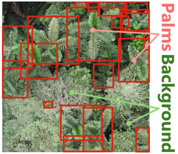


Figure: Bounding box annotations.

Figure: Georeferenced palm centers.

- Data collected from four reserves spanning wet to dry tropical forests.
- Captures high variation in palm species, density, and canopy structure.
- High-resolution orthomosaics created from thousands of UAV images.

The PRISM Framework

The PRISM Framework at a Glance

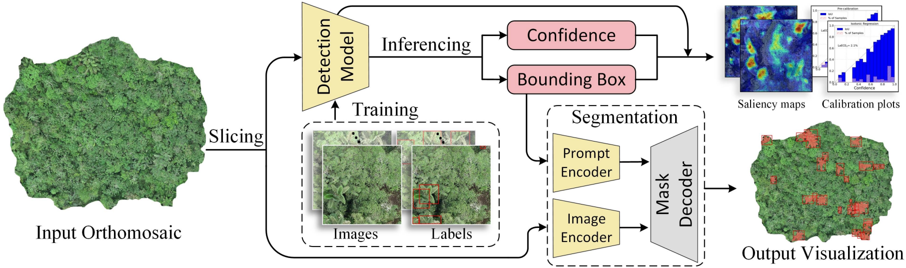


Figure: The PRISM Pipeline: From detection to segmentation and analysis [3].

Core Components:

- **Detection:** Fine-tuned models locate palms in orthomosaic patches.
- **Segmentation:** Detections are used as prompts for a zero-shot Segment Anything Model (SAM) to generate precise masks.
- **Mapping:** Outputs are georeferenced for landscape-scale analysis.
- **Interpretability:** Grad-CAM and calibration analysis enhance model reliability.

Detection and Segmentation Performance

Table: Detection model performance comparison.

Model	GFLOPS ↓	Params (M) ↓	FPS ↑	Precision ↑	Recall ↑	AP_{50} ↑	AP_{75} ↑	mAP ↑
DINO	1920.3	218.2	18.98 ± 0.95	0.7629 ± 0.0177	0.8494 ± 0.0071	0.8169 ± 0.0166	0.5455 ± 0.0150	0.5102 ± 0.0101
DDQ	1232.6	218.6	19.18 ± 0.96	0.7825 ± 0.0124	**0.8566 ± 0.0123**	0.8541 ± 0.0129	0.6354 ± 0.0137	0.5736 ± 0.0130
RT-DETR	222.5	65.5	151.49 ± 0.70	**0.8869 ± 0.0230**	0.7598 ± 0.0310	0.8416 ± 0.0181	0.6198 ± 0.0181	0.5769 ± 0.0145
YOLOv8	226.7	61.6	174.92 ± 0.86	0.8729 ± 0.0165	0.7997 ± 0.0203	0.8667 ± 0.0141	0.6777 ± 0.0137	0.6148 ± 0.0128
YOLOv9	**169.5**	53.2	114.96 ± 0.30	0.8763 ± 0.0176	0.7976 ± 0.0209	**0.8741 ± 0.0109**	0.6762 ± 0.0146	0.6162 ± 0.0122
YOLOv10	169.8	**31.6**	**177.04 ± 1.14**	0.8716 ± 0.0121	0.7968 ± 0.0089	0.8626 ± 0.0129	**0.6794 ± 0.0112**	**0.6173 ± 0.0090**
YOLO11	194.4	56.8	170.40 ± 0.95	0.8721 ± 0.0095	0.7896 ± 0.0127	0.8684 ± 0.0108	0.6677 ± 0.0180	0.6115 ± 0.0109

Key Findings:

- **YOLOv10 (Selected):** Best overall trade-off, achieving the highest mAP, AP_{75} and inference speed with the fewest parameters.
- **DDQ:** Highest recall, ideal when finding all possible instances is prioritized.
- **RT-DETR:** Highest precision, but misses more palms (lower recall).

Segmentation Performance: Comparing SAM Variants

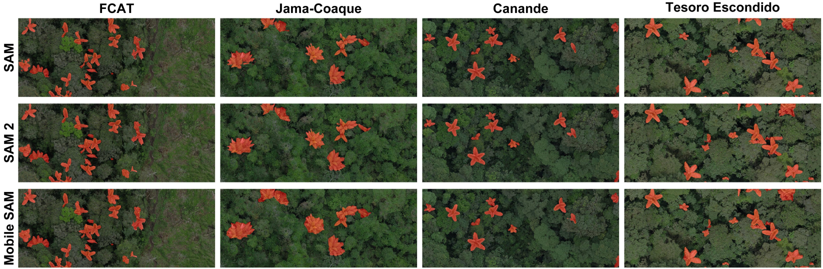


Figure: Visual comparison of SAM variants for zero-shot palm segmentation.

Key Findings

- We use the detector's bounding boxes as prompts for **zero-shot segmentation**.
- A comparison revealed distinct behaviors on our dataset:
  - **Original SAM:** Occasionally produces incomplete segments (under-segments).
  - **MobileSAM:** Tends to over-segment into non-palm areas.
  - **SAM 2 (Selected):** Provides the most balanced and accurate segmentation.

Visualizing What the Model "Sees" with Grad-CAM

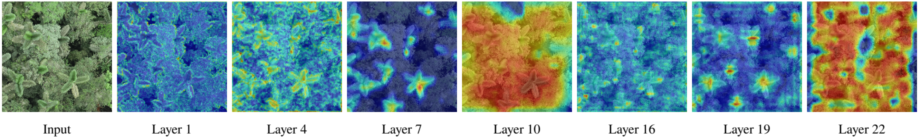


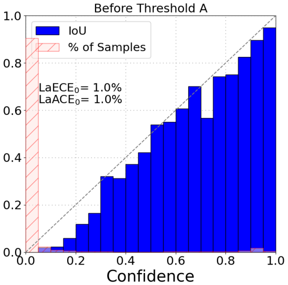
Figure: Hierarchical Feature Learning in YOLOv10 through Grad-CAM Visualizations.

Hierarchical Feature Learning

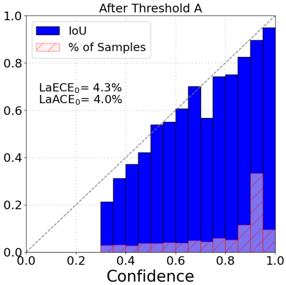
The analysis confirms the model learns a meaningful progression: early layers focus on low-level edges and textures; intermediate layers integrate spatial context; and deep layers exhibit focused activation over entire palm crowns.

Model Interpretability: A Step-by-Step Guide to Calibration

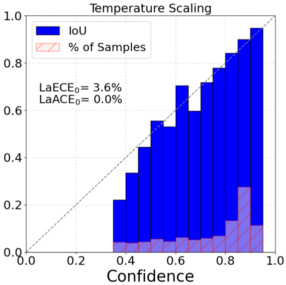
1. Initial Output



2. LRP Thresholding



3. Post-Hoc Calibration



The Problem

The uncalibrated model is unreliable; its confidence scores are poorly correlated with true accuracy (IoU).

The First Fix

LRP-based thresholding is first applied to prune the large number of unreliable, low-confidence predictions.

The Final Result

A post-hoc method (e.g., Temperature Scaling) is then applied to align confidence with accuracy.

Quantitative Analysis of Counting Performance

Table: Counting performance across four distinct ecological sites.

Site	Area (ha)	Counts	Pred2GT		GT2Pred	
			Ratio	Median (m)	Ratio	Median (m)
FCAT	21.62	471	0.9361	1.10	0.8854	0.77
Jama-Coaque	111.93	952	0.9348	1.50	0.8151	1.14
Canande	101.20	1,273	0.8956	0.82	0.7667	0.72
Tesoro Escondido	86.76	2,330	0.8981	1.09	0.9253	0.91

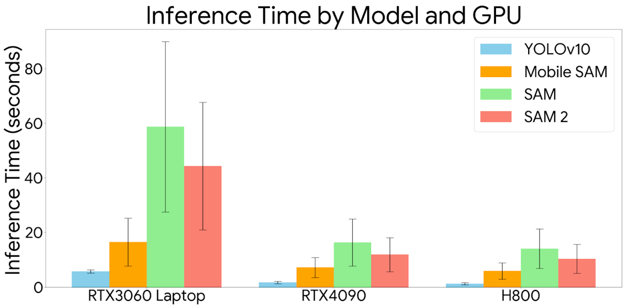
Key Metrics:

- **Pred2GT Ratio (Precision):** Proportion of predictions matched to a ground truth palm.
- **GT2Pred Ratio (Recall):** Proportion of ground truth palms matched by a prediction.

Summary of Findings:

- Precision is high across all sites, indicating the model generates few false positives.
- Recall is more variable, showing that detecting every true palm is harder and site-dependent.
- Sites like Tesoro Escondido show balanced performance, while Canandé reveals recall limitations (some palms are missed).

Feasibility for Real-Time Analysis



Key Findings:

- **Detection is Real-Time Ready:** YOLOv10 is fast enough (1.2–5.7s / image) for live processing on a UAV, even with mid-range hardware.
- **Segmentation is Costly:** Segmentation speed varies greatly, making it an optional step for time-critical missions.
- **Conclusion:** The core detection pipeline is efficient and stable, meeting the requirements for field deployment.

Spatial Distribution Modeling

Background: Measuring Spatial Point Patterns

Goal

To quantify if a spatial point pattern is clustered, random, or regular by comparing it against a model of Complete Spatial Randomness (CSR).

Ripley's G & F Functions

These are cumulative distribution functions (CDFs) that measure nearest-neighbor distances at a given distance radius d:

- $G(d)$: CDF of distances from each point to its *nearest neighbor in the pattern*. It quantifies internal clustering.
- $F(d)$: CDF of distances from *random locations* to the nearest point in the pattern. It quantifies empty space.

Formal Definitions

Function	Formulation
$G(d)$	$\frac{1}{N_o}\sum_{i=1}^{N_o}\mathbb{1}(\hat{d}_i < d)$
$F(d)$	$\frac{1}{N_r}\sum_{j=1}^{N_r}\mathbb{1}(\tilde{d}_j < d)$
$J(d)$	$\frac{1-G(d)}{1-F(d)}$

Here, N_O is the number of observed points and $\hat{d}_i$ is the distance from point i to its nearest neighbor. N_r is the number of random points and $\tilde{d}_j$ is the distance from random point j to the nearest observed point.

Modeling Palm Distributions

Our Goal

To simulate palm spatial patterns that match observed distributions and to understand the ecological drivers of reproduction (e.g., long-range dispersal vs. local clustering).

The Core Mechanism: A Hybrid Generative Process

The model simulates palm propagation by combining two key ecological processes, controlled by two interpretable parameters [4]:

- **Global Dispersal (Poisson):** With probability $(1-p)$, a new palm is placed randomly, representing animal-mediated or long-range dispersal.
- **Local Clustering (Gaussian):** With probability p, a new palm is placed near a parent, drawn from $\mathcal{N}(\mathbf{x}_{\text{parent}}, \sigma^2\mathbf{I})$, representing local seed drop.

Parameter Fitting

The optimal parameters (p^*, σ^*) are found by identifying the pair that generates simulated patterns whose spatial statistics (Ripley's G and F functions) most closely match those of the observed data.

The Poisson-Gaussian Algorithm: Implementation Details

Algorithm Pseudocode

Input: Candidate params $(\mathbf{p}, \boldsymbol{\sigma})$, Observed points X
Output: Optimal params (p^*, σ^*)

```
Initialize: d_min ← ∞
Pre-compute: Observed Ripley's stats G_obs, F_obs.
for each (p, σ) in grid do
    d_total ← 0
    for i = 1 to N simulations do
        1. Generate simulated set X̂_i via the
      Poisson-Gaussian process.
        2. Compute simulated stats G_sim, F_sim.
        3. Calculate discrepancy d_i.
        4. Add d_i to d_total.
    end for

    if d_total < d_min then
        Update d_min, p* ← p, σ* ← σ.
    end if
end for
return p*, σ*
```

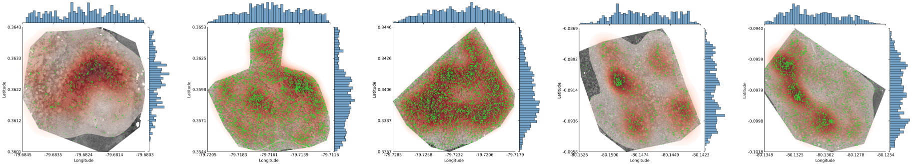
Discrepancy Metric

The discrepancy d_i for each simulation is the integrated absolute difference between observed and simulated Ripley's functions:

$$d_i = \int |\mathbf{g}_{obs} - \mathbf{g}_{sim}| + \int |\mathbf{f}_{obs} - \mathbf{f}_{sim}|$$

Optimization Process

A grid search is performed over the parameter space. The pair (p^*, σ^*) that minimizes the total discrepancy over all simulations is selected as the optimal fit.



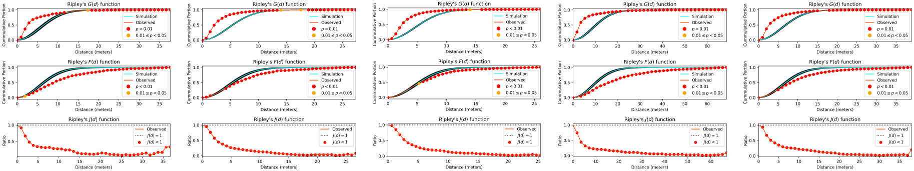
FCAT 1 | FCAT 2 | FCAT 3 | Jama-Coaque 1 | Jama-Coaque 2

Figure: Kernel Density Estimates (KDEs) of detected palm locations across five study sites, visually suggesting non-random clustering.

The Central Research Question

The spatial arrangement of palms appears clustered, but is this pattern statistically significant? We test the observed distributions against a null model of CSR.

Statistical Proof: Analysis with Ripley's Functions



FCAT 1 FCAT 2 FCAT 3 Jama-Coaque 1 Jama-Coaque 2

Figure: Ripley's function plots for each site. The observed pattern (red curve) is compared against the 95% confidence envelope of a CSR process (blue curve with shaded area).

Conclusion from the Analysis

The results confirm a **statistically significant** departure from randomness across all sites, with Ripley's functions revealing both dense internal clustering (G-function) and large empty spaces (F-function). This strong, non-random aggregation justifies our development of a more complex reproduction model.

Simulation Results: Replicating Observed Patterns

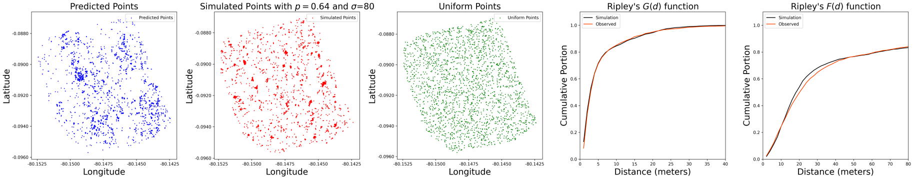


Figure: Visual and statistical comparison for the site Jama-Coaque 1. From left to right: PRISM prediction, model simulation, random distribution, G-, and F-function comparison.

Optimal Parameters

Fitted (p^*, σ^*) across sites.

Site	p^*	σ^*
FCAT 1	0.49	50
FCAT 2	0.52	70
FCAT 3	0.46	70
Jama-Coaque 1	0.64	80
Jama-Coaque 2	0.51	60

Key Findings

- The optimal parameters are highly consistent.
- This indicates a stable balance between local clustering (within a ~2-4 meter radius) and random, long-range dispersal.
- The strong alignment between simulated and observed Ripley's functions (right panels) validates the model's fidelity.

Conclusion

Conclusion

Summary of Contributions

- Created **PALMS**, a new, large-scale annotated dataset for palm detection in ecologically diverse tropical forests.
- Developed **PRISM**, an end-to-end framework for efficient palm detection, segmentation, and counting from UAV imagery.
- Introduced a simple, two-parameter **Poisson-Gaussian model** that successfully replicates the complex spatial dynamics of palm distribution, as validated by Ripley's functions.

Future Directions

- **Dataset Expansion:** A 1000 km^2 Amazonian region in Peru.
- **Model Enhancement:** Reduce supervision requirements and enable robust application on lower spatial-resolution imagery.
- **Deployment & Extension:** Real-time, on-device deployment (e.g., NVIDIA Jetson) and extension to species-level classification.

References I

[1] W. L. Eiserhardt et al. “Geographical ecology of the palms (Arecaceae): determinants of diversity and distributions across spatial scales”. In: *Ann Bot* 108.8 (2011), pp. 1391–1416.

[2] N. C. Pitman et al. “Distribution and abundance of tree species in swamp forests of Amazonian Ecuador”. In: *Ecography* 37.9 (2014), pp. 902–915.

[3] K. Cui et al. “Detection and Geographic Localization of Natural Objects in the Wild: A Case Study on Palms”. In: *Proc. IJCAI*. Accepted, to appear. 2025.

[4] K. Cui et al. “Efficient Localization and Spatial Distribution Modeling of Canopy Palms Using UAV Imagery”. In: *IEEE Transactions on Geoscience and Remote Sensing* (2025).

Thank You

Questions & Discussion

Jason Cui
Wake Forest University
ckn3.github.io