JOURNAL OF
EMERGING INVESTIGATORS

Article

Canopy complexity and plant morphology determine detectability in multi-perspective drone imagery

Dhruv Yadati[1], Ivy He[2]
[1] Monta Vista High School, Cupertino, California
[2] Brown University, Providence, Rhode Island

SUMMARY

Accurate identification of plant species in drone imagery is important for vegetation monitoring, but plant morphology, including branching, leaf orientation, and vertical layering, influences how reliably deep-learning models can detect and classify plants in these images. Canopy complexity, in particular, can obscure diagnostic features, making some species more challenging to identify than others. Single-perspective surveys often miss plants that are partially hidden beneath dense or layered canopies. In this study, we evaluated whether multi-perspective drone imagery improves classification accuracy compared to aerial-only surveys by testing the role of the canopy structure of crops. We hypothesized that species with more complex canopies would show greater improvements in classification accuracy when analyzed with multi-perspective drone imagery. We analyzed approximately 1,250 images from five agricultural crop species that span a gradient of morphological complexity using a deep learning pipeline that combined RetinaNet and ResNet-50 models. We found that overall classification accuracy increased from 35.73% with aerial imagery to 52.93% with multi-perspective imagery. Species with structurally complex canopies showed the largest gains in accuracy, whereas low-complexity crops exhibited only marginal improvement. By linking plant morphology to model performance, our study shows that multi-perspective drone surveys are especially valuable for identifying species with complex canopies. These findings can guide field operators in designing effective drone-based monitoring strategies, particularly for invasive plant species with dense or multilayered canopies that are often missed in traditional aerial surveys.

INTRODUCTION

The United States harbors about 4,500 invasive plant species, and non-native plants make up roughly 8 to 47% of state floras (1). About 15% of these species cause major ecological or economic impacts, disrupting ecosystems and food webs by displacing native plants and competing for water, nutrients, and sunlight (1, 2). Certain invaders with complex structural traits can compound the ecological impacts and make management of the invasive species more difficult. In Hawaiʻi, nearly half of the native forests are invaded (3). One of the most problematic invaders is Moluccan albizia (*Falcataria falcata*). This tree species grows quickly and builds a broad and highly branched canopy, the upper layer of foliage formed by the crowns of leaves, branches, and stems, which shades native understory plants. Because the canopy is the part of a plant that faces the sky, its density and vertical arrangement determine how much of the plant an aerial sensor can see. Its brittle wood creates hazards that are harder to manage for detection and eradication (4, 5). The overlapping crown structure of Albizia covers the canopies and makes it difficult for aerial surveys to detect it. Similarly, Japanese knotweed (*Fallopia japonica*) and Brazilian pepper (*Schinus terebinthifolia*) spread through dense, layered canopies and make detection difficult (6, 7). Across the United States, invasive plants like these contribute to nearly $120 billion in annual losses, while delayed management often leads to costs many times higher than early intervention (8, 9). These challenges highlight the need for methods that account for canopy structure in detection systems.

Plant architecture provides a biological framework for understanding this problem. Traits such as branching patterns, leaf orientation, and canopy layering influence how species appear from different viewing angles (10, 11). Plant imagery, including photographs or sensor data collected remotely from aircraft, drones, or satellites and interpreted to identify species and map their distribution, has become central to vegetation monitoring because it can cover large areas repeatedly without disturbing plants or requiring ground surveyors. Upright and uniform canopies can often be recognized from top-down imagery alone, while species with heterogeneous or multilayered canopies require additional perspectives to capture diagnostic features. The importance of vertical structure is further underscored by studies showing that leaf spectral traits vary with canopy layer and that classification accuracy can improve when combining trait information across canopy strata (12, 13). This dimension of biology is central to detection but has received relatively little attention compared to algorithmic advances. Remote sensing measures objects from a distance by recording the light or other energy they reflect or emit, rather than by sampling them directly in the field. Studies in this field typically evaluate model performance based on the proportion of plants correctly located and classified against verified reference annotations. Most efforts to improve this performance focus on modifying algorithms rather than examining why certain plants are consistently more difficult to detect or classify (14, 15). Expanding the focus to include biological structure may therefore provide important insights into detection outcomes, particularly through approaches designed to capture self-occluded plant regions hidden behind the plant's own leaves and branches rather than neighboring vegetation (16).

Remote sensing technologies already play a major role in vegetation monitoring. Drones equipped with red-green-blue (RGB), infrared, or light detection and ranging (lidar) sensors can rapidly survey large landscapes at fine resolution (17, 18). Close-range flights with quadcopters and

https://doi.org/10.59720/25-277

hexacopters capture lateral views that complement traditional aerial images (19). While lidar can penetrate canopies, it is costly and limited in scope (18, 20). Multi-perspective drone imagery, by comparison, offers a lower-cost and flexible way to capture structural information and has been shown to approximate canopy parameters with accuracy comparable to lidar mounted on uncrewed aerial vehicles (UAVs) (21).

Machine learning (ML) further strengthens these tools by improving image analysis. Traditional models such as Support Vector Machines and Random Forests have been widely used for vegetation mapping, while convolutional neural networks and object detection frameworks such as You Only Look Once (YOLO) are increasingly common for extracting complex features (22, 23, 24). Deep learning has been applied in both agriculture and invasive plant detection, with lightweight models such as IPMCNet and MAR-YOLOv9 developed for real-time use (25, 26). Despite these advances, classification accuracy still varies across species, particularly those with complex canopy structures that obscure diagnostic features when only a single perspective is available (14).

Existing work on multi-perspective UAV imagery suggests that plant structure shapes how well imaging performs but has not tested that relationship directly. Deep learning can integrate multiple viewing angles within a single neural network, whether applied to coastal terrain or to digital twin representations of foliage (27, 28). Furthermore, adding viewpoints improves phenotypic trait extraction from soybean canopies enough to permit accurate plant height estimation from 3D reconstruction (29). The most suggestive result comes from forest surveys, where UAV-derived Structure-from-Motion point clouds track lidar measurements closely for homogeneous stands but degrade for heterogeneous ones (30). That degradation implies that structural complexity itself limits what a single perspective can resolve, yet no study has tested this implication by treating canopy complexity as the variable of interest and measuring how much benefit each species draws from additional viewing angles. Our study addresses this gap by testing whether crops with greater structural heterogeneity derive disproportionate benefit from multi-perspective imaging.

This study used five crop species (rice, wheat, maize, sugarcane, and jute) as model systems to investigate how canopy complexity influences classification accuracy. We hypothesized that species with structurally more complex canopies would show greater improvements in classification accuracy when analyzed with multi-perspective drone imagery. Crops provide a tractable system because of the availability of high-quality labeled imagery across growth stages (31) and because their canopy architectures are well characterized and can be compared under controlled structural conditions. By comparing performance across species with different canopy architectures, we tested whether plant structural traits predict detection outcomes. We found that multi-perspective imagery raised overall classification accuracy from 35.73% to 52.93%, with the size of that gain scaling with canopy complexity, from +4.2% for rice, the structurally simplest crop, to +38.1% for sugarcane, the most complex. By incorporating biological structure into ML classification, this study provides a framework for designing detection strategies that reflect plant morphology.

RESULTS

The goal of this study was to determine whether the benefit of multi-perspective drone imagery depends on the structural complexity of a plant's canopy. We hypothesized that species with more complex canopies would show greater improvements in classification accuracy under multi-perspective imaging than under top-down aerial imaging alone. To test this hypothesis, we compared classification accuracy across five crop species spanning a gradient of

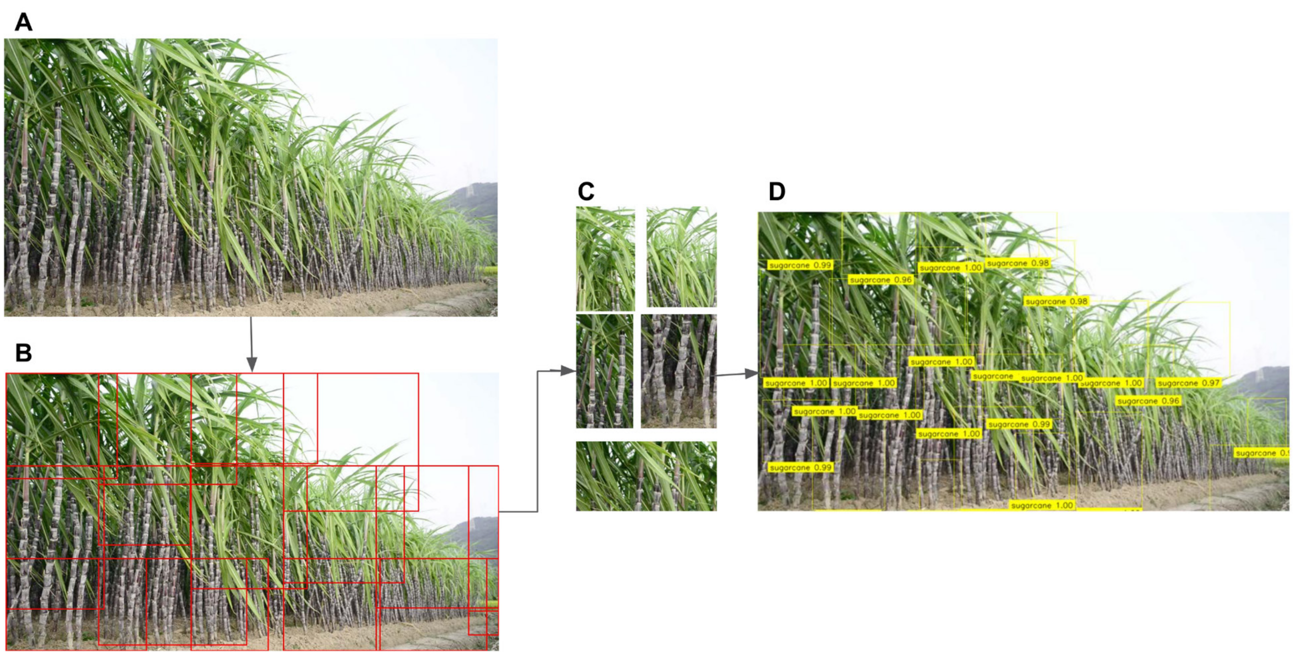

Figure 1. Pipeline for object detection and crop classification using a sugarcane image. (A) Original image. (B) Detection of candidate crop regions using a RetinaNet model with bounding boxes. (C) Cropped image tiles extracted from bounding boxes. (D) Classification and labeling of crops using a ResNet-50 model, with predicted labels overlaid on the original image.

https://doi.org/10.59720/25-277

canopy architectures. The ML pipeline classified five crop classes using two models in sequence: (a) Object Detection with RetinaNet to locate bounding boxes of plants in images, and (b) Object Classification with ResNet-50 to categorize and label the boxed plants (**Figure 1**). We measured performance as classification accuracy, the proportion of object-level predictions that assigned the correct crop label. We performed predictions on 50 new plant images, approximately 10 from each of the five crop classes.

Overall accuracy and score confidence

We first examined whether multi-perspective imagery improved crop classification compared with top-down imagery alone. We set the RetinaNet threshold very low during detection so that candidate boxes were not discarded before the ResNet classifier could evaluate them. At inference, after removing small, noisy boxes with a minimum width and height of 80 pixels (px) and a minimum area of 6,400 px, a subset of object-level predictions was retained for each dataset. Overall accuracy was higher for multi-perspective imagery, which pairs top-down aerial views with oblique and side-angle views that expose vertical canopy structure, than for aerial-only imagery, which uses top-down views alone (52.93% vs. 35.73%). This gain of 17.2 percentage points was consistent with our hypothesis that additional viewing angles supply structural information a classifier can use, and it was larger than we anticipated for a change in training imagery alone. Accuracy increased monotonically with classifier score, and high-confidence predictions (score ≥ 0.9) reached 71.2% accuracy for the multi-perspective-trained model and 42.4% for the aerial-trained model (**Figure 2**). That the ordering held at every confidence level, rather than only in aggregate,

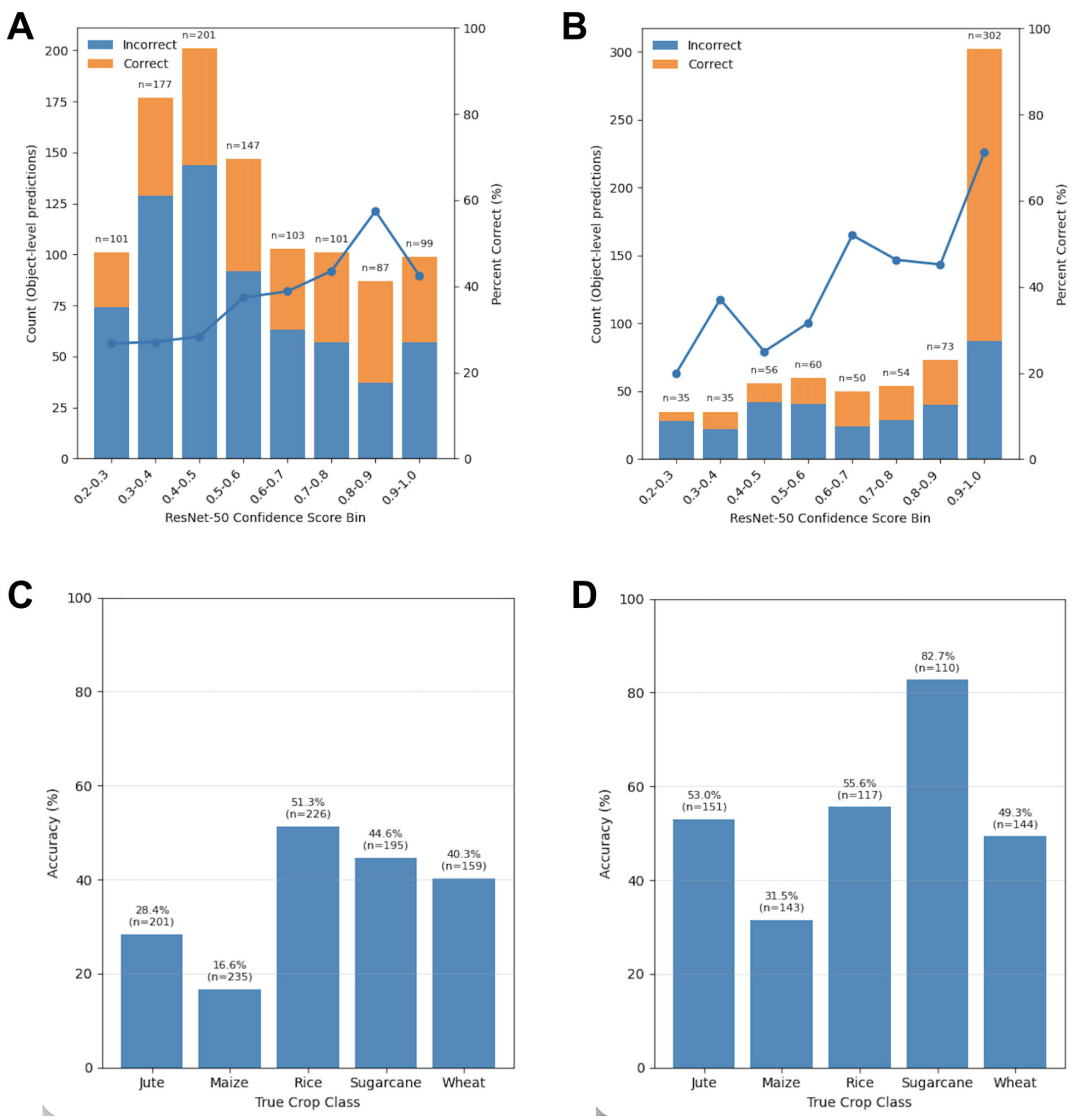

Figure 2. Score-binned prediction outcomes and per-class accuracy on the held-out test set. The left column (A, C) shows the aerial-trained model, and the right column (B, D) shows the multi-perspective-trained model. (A–B) Stacked bars show incorrect and correct predictions within each ResNet-50 confidence-score bin. The overlaid line shows percent correct, and n indicates the total number of predictions per bin. (C–D) Bars show classification accuracy by true crop class, with values labeled on the bars and n indicating sample size. Both models were evaluated on the same independent test set of 50 images, including 10 images per crop, after applying the 80 × 80 px object-size filter. Multi-perspective training improved accuracy overall, with the largest gains observed for sugarcane and jute.

JOURNAL OF
EMERGING INVESTIGATORS

https://doi.org/10.59720/25-277

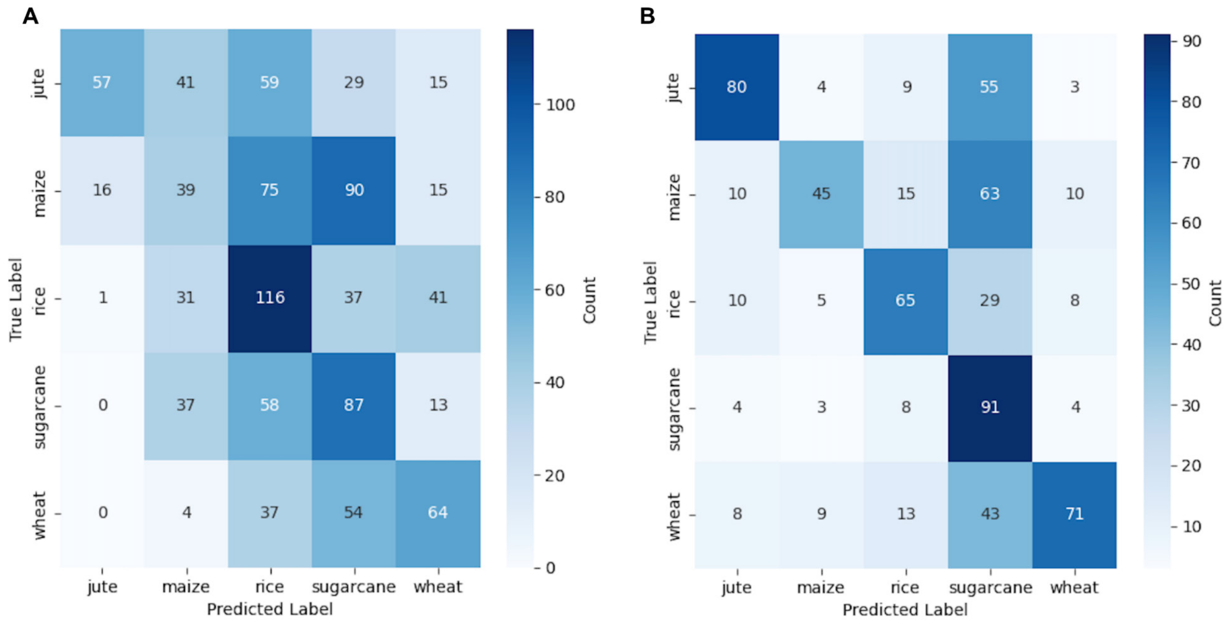

Figure 3. Confusion matrices for predictions on the held-out test set images. Confusion matrices summarize object-level predictions generated by (A) the aerial-trained model and (B) the multi-perspective-trained model when applied to the same independent test set after applying the minimum pixel size filter. Rows indicate true labels and columns indicate predicted labels. The multi-perspective-trained model shows a stronger diagonal (overall accuracy 52.93% vs. 35.73%), with the largest gains for high-complexity crops such as sugarcane (82.7% vs. 44.6%) and jute (53.0% vs. 28.4%). Common confusion patterns include maize→sugarcane and wheat→sugarcane.

indicates the multi-perspective model was better calibrated and not merely more accurate on average, which matters in practice because operators filter predictions by confidence.

Crop-specific performance

Performance varied by crop class and was consistent with the canopy complexity hypothesis. Under multi-perspective imaging, sugarcane achieved the highest accuracy (82.7%), followed by rice (55.6%), jute (53.0%), wheat (49.3%), and maize (31.5%). Under aerial imaging, rice performed best (51.3%), followed by sugarcane (44.6%), wheat (40.3%), jute (28.4%), and maize (16.6%) (**Figure 2**). Sugarcane and jute, the two crops with the largest gains, possess tall, vertically layered, and heterogeneous canopies that expose different diagnostic features across viewing angles, which is the pattern the hypothesis predicts. Maize, however, ranked last under both imaging strategies despite its intermediate complexity score, which we did not expect. Its narrow, steeply inclined leaves may present little distinctive texture from any single angle, and maize was the class most frequently confused with sugarcane (**Figure 3**). This finding suggests that complexity alone does not guarantee distinctiveness: a canopy must be both structurally rich and structurally unlike its neighbors to be separable. Although rice achieved relatively high absolute accuracy under multi-perspective imaging, it exhibited the smallest gain because its aerial-only accuracy was already relatively high. Confusion matrices of the object-level predictions from the 50 held-out test images illustrate the distribution of errors for each dataset and show a clearer and more consistent diagonal across multiple crop classes under multi-perspective imaging, most notably for the structurally complex crops sugarcane and jute, while aerial imaging exhibits a pronounced diagonal primarily for rice (**Figure 3**).

Relationship with canopy complexity

To connect biological structure with model behavior, we used a rubric from plant architecture literature to score canopy complexity for each crop and then compared these scores with the benefit of multi-perspective imaging. The rubric assigned rice and wheat to low complexity (1), because their canopies are vertically uniform when viewed from drone altitudes and wheat tillering occurs near the plant base where it is not visually accessible from above; maize to medium complexity (2), because of its variable leaf angles and vertically distributed organs, which become more apparent from oblique views; and jute and sugarcane to high complexity (3), because of their tall, vertically layered, and structurally heterogeneous canopies (**Figure 4**). The accuracy gain of multi-perspective over aerial imaging increased with canopy complexity, with low-complexity crops such as rice and wheat showing the smallest improvements (+4.2% and +9.1%), intermediate gains for maize (+14.9%), and the largest gains for the structurally complex crops jute (+24.6%) and sugarcane (+38.1%) (**Figure 5**). To measure the correlation between canopy complexity score and accuracy improvement, we used Spearman's rank correlation (ρ), a non-parametric test suitable for ordinal data that does not assume a linear relationship; a significant positive ρ indicates that higher-complexity crops tend to show greater improvement from multi-perspective imagery. The test returned $\rho = 0.949$ with $p = 0.014$, a significant positive relationship between the two variables. A one-way ANOVA, which tests whether mean accuracy improvements differ significantly across the low, medium, and high complexity groups and therefore whether the observed pattern could arise by chance, did not yield statistically significant differences ($F = 6.09$, $p = 0.141$), likely reflecting the small number of species ($n = 5$) (**Figure 5**).

https://doi.org/10.59720/25-277

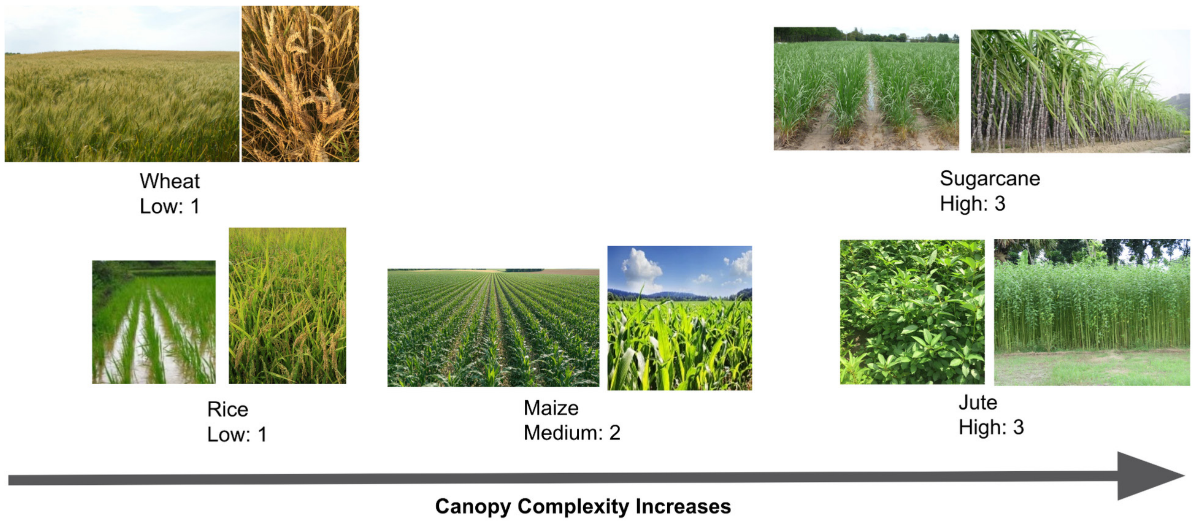

Figure 4. Ordinal canopy complexity categories for crop species. Representative examples of drone-observable canopy structure used to assign crops to low (1), medium (2), or high (3) complexity classes based on vertical layering, structural heterogeneity, and viewing-angle dependence. Complexity scores are ordinal and reflect relative differences in observable canopy architecture. Rice and wheat were assigned low complexity (1), maize medium complexity (2), and jute and sugarcane high complexity (3).

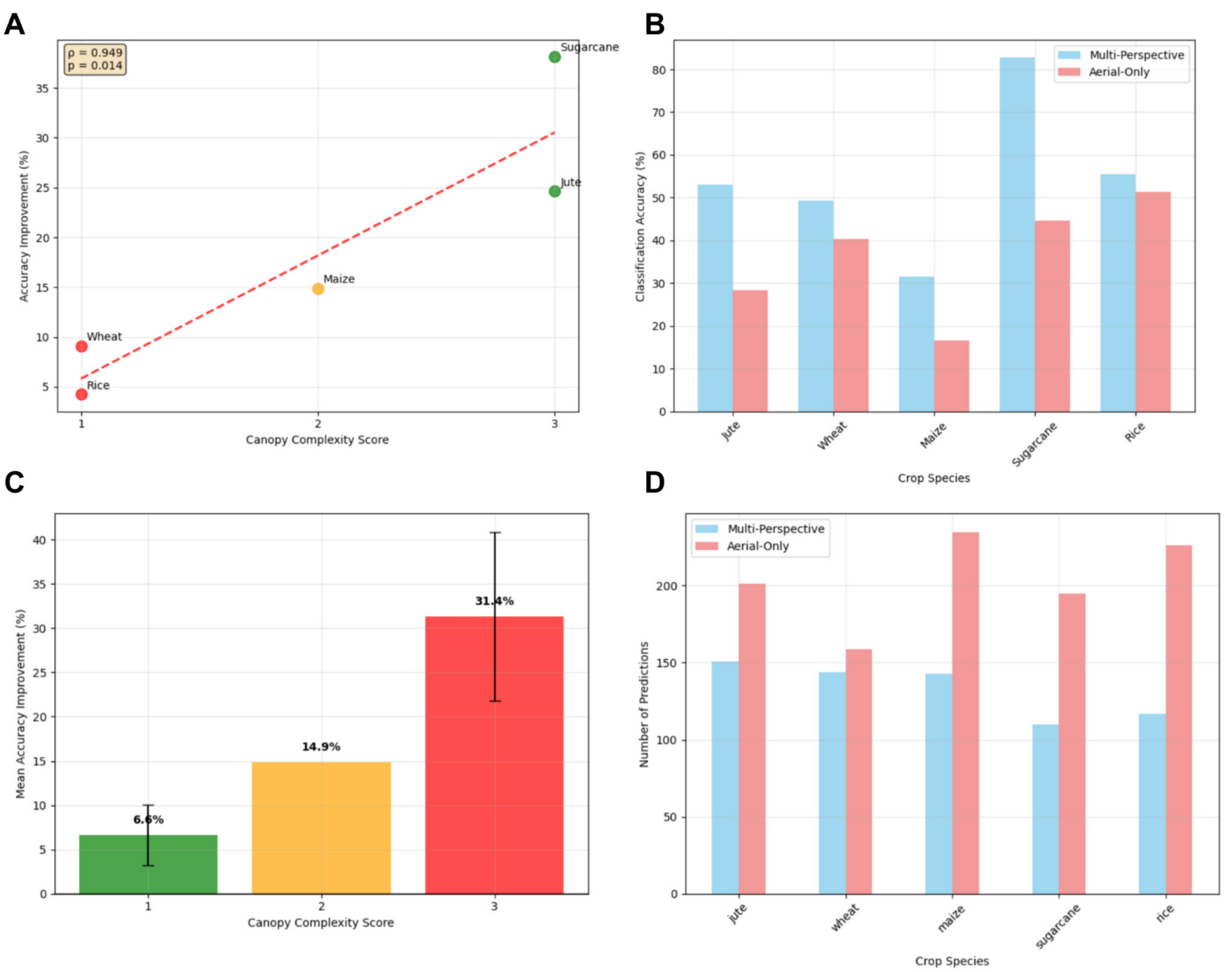

Figure 5. Relationship between canopy complexity and classification performance on the test set predictions. (A) Significant positive correlation between canopy complexity score and accuracy improvement from multi-perspective imaging relative to aerial imaging (Spearman $\rho = 0.949$, $p = 0.014$). (B) Classification accuracy by imaging method; multi-perspective outperforms aerial for all crops, with the largest gains in high-complexity species (sugarcane, jute). (C) Mean improvement by complexity group: low (6.6%), medium (14.9%), high (31.4%). (D) Number of object-level predictions generated per crop and imaging method from the same test images after filtering.

JOURNAL OF
EMERGING INVESTIGATORS
https://doi.org/10.59720/25-277

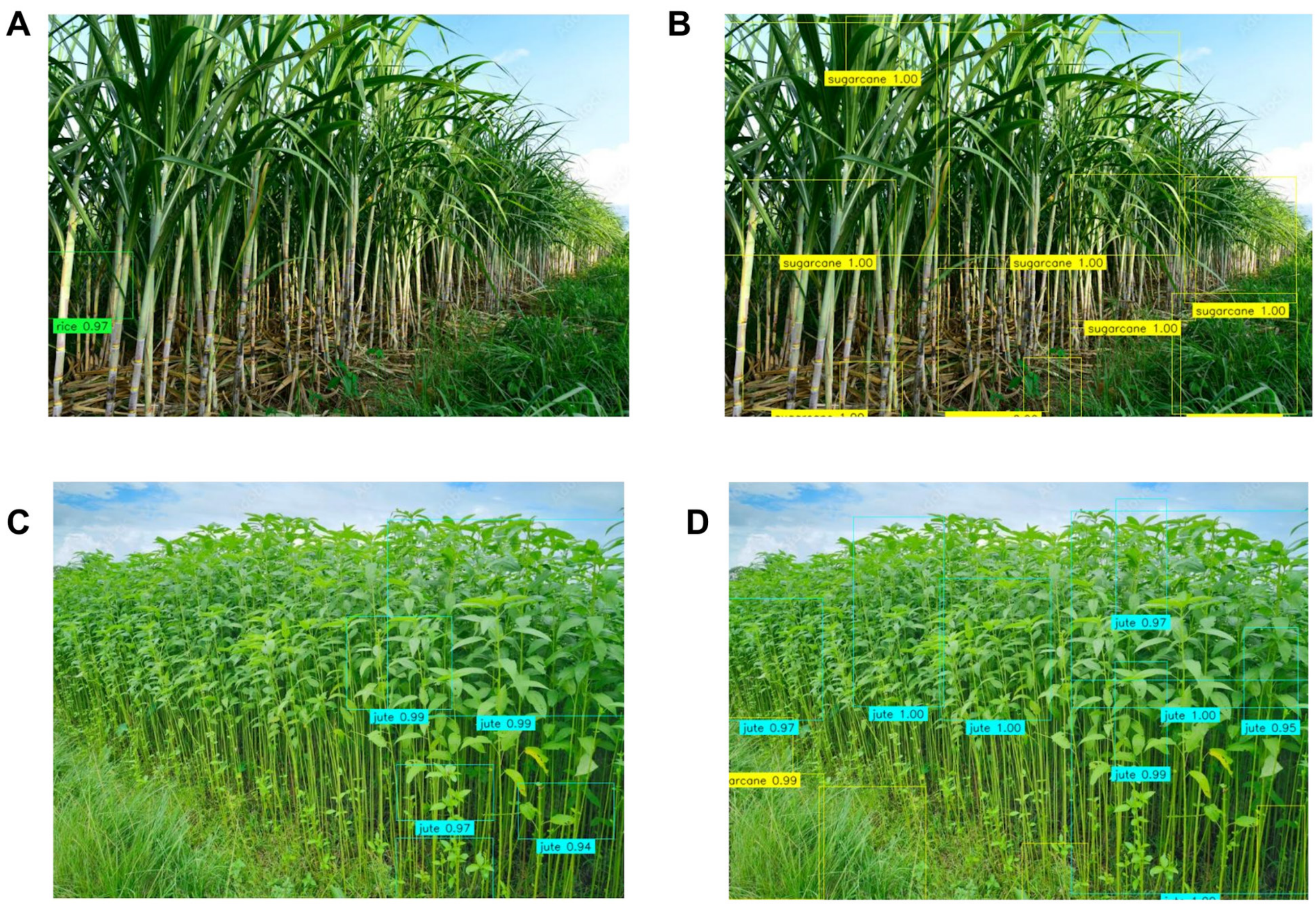

Figure 6. Predictions on high-complexity crops (sugarcane and jute) show that high-complexity crops with vertically layered canopies benefit substantially for classification accuracy from multi-perspective training data. (A) The aerial-trained model produces one incorrect classification (rice) on a sugarcane test image. (B) The multi-perspective-trained model correctly identifies multiple sugarcane regions with high confidence on the same image as A. (C) The aerial-trained model identifies a few jute regions correctly. (D) The multi-perspective-trained model detects more jute regions with correct labels on the same image as C.

Qualitative analysis of detections

We performed a qualitative assessment of detection patterns to complement the quantitative metrics and to test whether the accuracy differences arose through the mechanism we proposed. If additional viewing angles help by revealing diagnostic features hidden in aerial views, then the improvement should appear as visibly cleaner detections on structurally complex crops specifically, rather than as a uniform shift across all classes. For structurally complex crops such as sugarcane and jute, multi-perspective images produced cleaner and more spatially consistent detections, while aerial-trained models frequently misclassified regions or detected fewer true positives. In contrast, low-complexity crops such as wheat showed a different pattern, with multi-perspective training increasing detection recall but also introducing additional false positives, reflecting a recall–precision trade-off when additional viewing angles provide limited new structural information (**Figures 6 and 7**). Finally, imposing the 80 × 80 px minimum box filter reduced noisy, small boxes and clarified true positives in both models. This filtering reduced spurious rice detections in the aerial dataset and improved jute detection in the multi-perspective dataset (**Figure 8**).

DISCUSSION

This study asked whether the benefit of multi-perspective drone imagery depends on plant structure, using five crop species spanning a gradient of canopy architecture, from the vertically uniform canopies of rice and wheat to the tall, layered canopies of jute and sugarcane, to ask whether structurally complex species gain more from additional viewing angles than simple ones do. The results supported the central hypothesis that species with more structurally complex canopies derive greater benefit from multi-perspective drone imagery. Crops with high canopy complexity, particularly sugarcane and jute, exhibited the largest improvements in classification accuracy when we applied multi-perspective views. Maize, which exhibits intermediate canopy complexity, showed moderate improvement. In contrast, low-complexity crops such as rice and wheat exhibited relatively small gains from multi-perspective imagery; for example, rice achieved high absolute accuracy under both imaging strategies, but additional viewing angles provided little new structural information and resulted in only marginal improvement.

Two design choices in the pipeline were critical for obtaining these results. First, setting the RetinaNet detection score threshold very low during the detection stage prevented

https://doi.org/10.59720/25-277

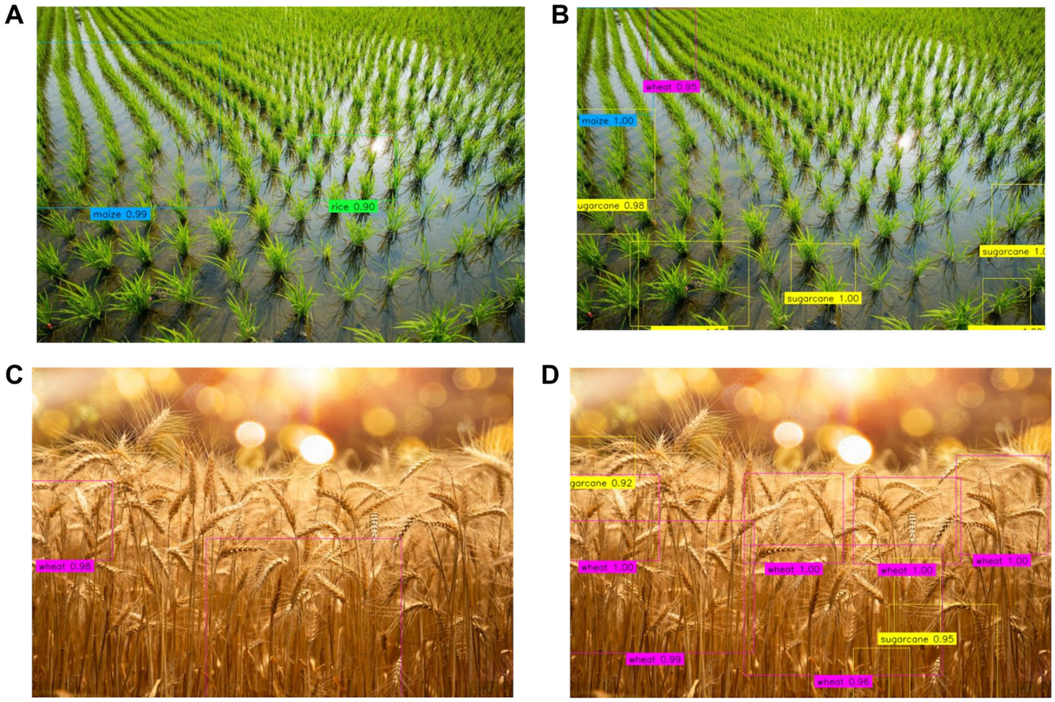

Figure 7. Predictions on low complexity crops (rice and wheat) show recall-precision trade-offs. (A) The aerial-trained model identifies one correct (rice) and one incorrect (maize) classification on rice crop test image, while (B) the multi-perspective-trained model recalls more boxes and produces multiple incorrect misclassifications on the same rice image as A. (C) Aerial-trained model detects two regions of wheat field correctly. (D) The multi-perspective-trained model detects more wheat regions (6 correct vs. 2) on the same wheat image as in C but also introduces false positives (2 sugarcane). This pattern suggests multi-perspective training increases detection recall but may reduce precision for low-complexity crops where additional viewing angles provide less discriminative information.

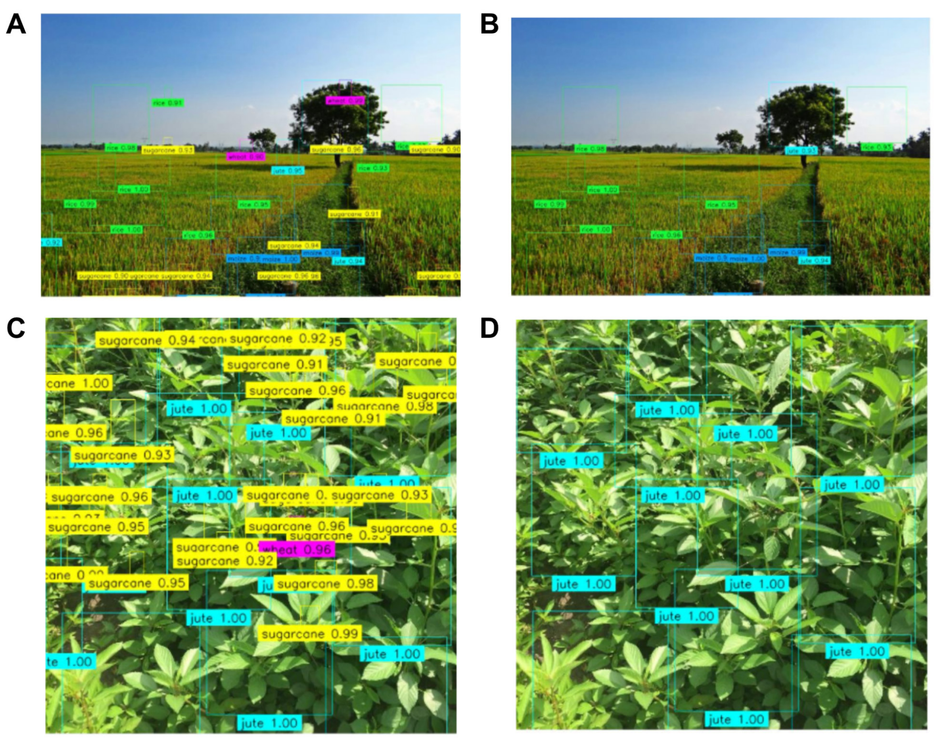

Figure 8. Effect of the 80 × 80 px minimum box filter on predictions. The left column shows predictions before filtering and the right column shows the same images after filtering. (A, B) Rice field processed by the aerial-trained model, where filtering removes spurious small detections. (C, D) Jute canopy processed by the multi-perspective-trained model, where filtering removes noisy sugarcane and wheat predictions, leaving cleaner jute detections. The 80 × 80 px minimum box size threshold reduces false positives while preserving true crop regions.

https://doi.org/10.59720/25-277

the loss of true positives before classification. When we set the threshold as high as 0.4, the pipeline discarded many boxes early and the classifier never had the opportunity to evaluate them. Separating detection (i.e., "where are the plants?") from classification (i.e., "what species are the plants?") and filtering only at the classification stage therefore matters more than the choice of either model. Second, applying a strict size filter at prediction time reduced the influence of small, noisy boxes, which often inflated false positives even when their confidence scores were above 0.9. After filtering to a minimum of 80 × 80 px, the remaining boxes represented larger, more informative patches, and the outputs of both models became easier to interpret. The difference was most apparent for rice and jute, where filtering removed spurious small detections without cost to true crop regions.

The biological findings indicate that canopy complexity strongly influences detectability. The contrast between the two statistical tests is itself informative: complexity correlated with benefit as a rank order but did not separate the crops into distinct groups. This is what a continuous relationship looks like when it is tested with only five species, the underlying trend is continuous, and the discrete groups impose boundaries the data do not, so with only five species the ANOVA has limited power to resolve group-level differences. Together, these results suggest that structural complexity exerts a continuous influence on classification performance rather than producing sharply separated categories. Complex layered canopies revealed different diagnostic features from different viewing angles, so side-view information complemented top-down texture and color cues. Practically, these results suggest that multi-perspective flight plans are most valuable when target taxa are architecturally complex or partially occluded, while top-down surveys may suffice for low-complexity canopies. Compared with lidar, which measures canopy structure directly by timing laser pulses reflected from vegetation surfaces and can therefore resolve height even beneath a canopy, multi-perspective RGB drone imagery infers structure indirectly from the way a plant's appearance changes across viewing angles. That indirect approach is markedly cheaper and requires no specialized sensor, and it was sufficient to improve performance for complex canopies in this study; lidar would likely remain advantageous where precise vertical structure or below-canopy penetration is required. Compared with satellite imagery, drones offer finer spatial resolution and flexible viewing geometry, which may be critical for identifying species with small diagnostic features or strong self-occlusion.

This study has a few limitations to consider. In this study, we worked with five agricultural crop species as controlled test systems to explore how canopy architecture influences classification accuracy. Using crops allowed us to take advantage of consistent imaging conditions and well-documented reference datasets to clearly evaluate the effects of structural complexity. At the same time, agricultural systems are more uniform than natural landscapes and therefore cannot fully capture the variability, patchiness, and mixed species interactions that are common in real-world ecosystems. Invasive plants often occur in mixed stands, grow in irregular patches, and experience greater occlusion, which makes them even harder to detect. Field deployment would introduce additional challenges not captured in our controlled dataset, including species mixing within single images, irregular patch boundaries, partial occlusion by other vegetation, and variable growth stages. We selected these crops rather than invasive species for three methodological reasons. First, the crops dataset (31) provides standardized, high-quality labeled imagery with verified ground-truth annotations, which is a prerequisite for training deep learning models and is rarely available for invasive species at comparable scale. Second, agricultural crops exhibit well-characterized morphological variation spanning a gradient of canopy complexity, from low-complexity cereals to high-complexity tall grasses, enabling controlled hypothesis testing about the relationship between plant architecture and imaging approach. Third, crop species allowed systematic isolation of structural variables, whereas natural invasive populations confound morphological complexity with environmental heterogeneity, species mixing, and irregular patch distribution. Because the source dataset compiled images across varied field conditions, environmental variability was not explicitly controlled. While this introduces potential noise, it also reflects realistic deployment scenarios for drone-based monitoring, and although systematic variation in lighting or image quality across crop types could confound results, no such bias was observed during manual image review. The test set also contained only 50 images (10 per species), so per-class accuracies should be interpreted as preliminary estimates rather than precise benchmarks. We recommend that future validation efforts include: (a) field trials with manually verified ground-truth from GPS-tagged locations; (b) testing across multiple growth stages and seasons; and (c) gradual expansion from monoculture plots to mixed-species environments.

Results from this study should be interpreted as establishing proof-of-concept for the relationship between canopy complexity and imaging strategy rather than operational field accuracy. Future work should test challenging species such as Moluccan albizia, Japanese knotweed, and Brazilian pepper in field conditions where occlusions and complex canopy layers are common (1–7). Furthermore, we did not explicitly model environmental factors such as sun angle, wind, and growth stage. These factors could generally influence how plants appear in images and affect detection and classification results as well. Incorporating these factors can make the model more robust to represent realistic conditions of forest imagery.

Finally, we reduced noise by filtering small boxes. Future improvements could use learned non-maximum suppression designed for multi-view images or fuse views in 3D. Despite these limitations, our findings show that multi-perspective drone imagery provides a biologically informed way to improve ML-based plant detection when canopy structures are complex and layered. The results also showed that traditional top-down aerial imagery remains effective for plants with simple, uniform canopies, where additional perspectives do not add meaningful information.

MATERIALS AND METHODS

Study design and datasets

We conducted a comparative study using two image acquisition strategies to evaluate how different viewing perspectives influence plant classification performance. The aerial dataset used straight-down images that mirror standard drone surveys. The multi-perspective dataset paired those top-down views with lower-altitude oblique shots that reveal

https://doi.org/10.59720/25-277

lateral canopy features. Images covered five crops, namely rice (*Oryza sativa*), wheat (*Triticum aestivum*), maize (*Zea mays*), jute (*Corchorus spp.*), and sugarcane (*Saccharum officinarum*), sourced from a public dataset curated for agricultural classification (31). Since high-quality labeled imagery from natural ecosystems, including invasive species, is often limited, we used crops as model systems to test the canopy-complexity hypothesis under controlled conditions. This approach allowed us to isolate the role of plant structure in detection, while still generating insights that can inform monitoring in more complex real-world settings.

We used approximately 1,250 images in total (~250 per crop), of which 85% were used for training and 15% for validation. The aerial dataset was curated to include only top-down (nadir) views representative of standard drone surveys conducted at altitudes of 50-100 feet. The multi-perspective dataset retained these top-down images while incorporating additional oblique and side-angle views where vertical canopy structures such as sugarcane stalks and jute stems are visible. This distinction tests whether training on viewing-angle diversity improves classification of structurally complex crops. Environmental variability (lighting conditions, wind-induced motion blur, shadows) was not explicitly controlled during image acquisition, as the source dataset compiled images across varied field conditions. Images were reviewed manually to check for systematic variation in lighting or image quality across crop types. For evaluation, we used an independent test set of 50 images (10 per crop) with no overlap with the training or validation data. These test images generated multiple object-level predictions after detection and filtering, and all reported figures and confusion matrices summarize statistics computed from these predictions rather than from raw image counts (**Figure 3**).

Model architecture and training

RetinaNet produced predicted boxes from anchor grids and used focal loss to handle class imbalance (24). We then cropped those boxes and classified them with a ResNet-50 convolutional neural network (32). We used transfer learning by initializing from ImageNet weights and fine-tuning on the crop images. Both RetinaNet and the ResNet classifier were implemented with the DeepForest framework and its utilities (33-35). The RetinaNet detection model was trained using focal loss, which down-weights well-classified examples to focus learning on difficult-to-detect objects, which is particularly useful when most image regions contain background rather than crops. The ResNet-50 classification model was trained using cross-entropy loss, which measures the difference between predicted class probabilities and true labels, penalizing confident incorrect predictions more heavily. Hyperparameters were selected by monitoring validation loss. The classification model (ResNet-50) was trained for 10 epochs. The detection model (RetinaNet) was trained for 150 epochs with a learning rate of 0.0001 and batch size of 4. Both models were initialized with ImageNet-pretrained weights. At inference, non-maximum suppression used an intersection-over-union threshold of 0.15, and images were processed in 400 × 400 px patches with 5% overlap.

Annotation procedures and bias controls

Ground-truth crop labels and bounding box annotations were provided by the source dataset (31). Annotations were reviewed to verify label consistency and alignment with visible crop features. Images with full-frame bounding boxes (224 × 224 px) were excluded from detection training but retained for classification training, as they do not provide localization information. Both models employed an 85/15 train/validation split with a fixed random seed (seed = 42) to ensure reproducibility. For the classification model (ResNet-50), cropped plant images were randomly partitioned using PyTorch's torch.randperm() with subset indexing to prevent data leakage. For the detection model (RetinaNet), we used a stratified split at the image level to ensure all crop classes were represented in the validation set. Data augmentation included random horizontal flips and color jitter (brightness, contrast, saturation, hue).

Detection thresholding and post-processing

During detection, the RetinaNet score threshold was set to 0.01 so that candidate boxes were not discarded before classification. After classification, predictions were analyzed across score buckets, and high-confidence visualizations used a minimum classifier score of 0.9. To reduce noise from very small boxes, a size filter requiring a minimum width and height of 80 px and a minimum area of 6,400 px was applied to all quantitative analyses. This threshold was derived from the imaging geometry. Given a typical ground-sample distance (GSD) of 1–3 cm/px for drone-based imagery, where GSD is the ground distance represented by a single image pixel, and diagnostic canopy features on the order of 10–15 cm, resolving a feature requires at least 8 px across it; applying a safety factor of 1.5 yields a minimum box side of approximately 80 px. Predictions dominated by sky or background regions were additionally removed using a simple color-based filter to exclude bounding boxes containing predominantly blue or high-brightness pixels, reducing false positives unrelated to vegetation. After filtering, the multi-perspective dataset retained 665 predictions and the aerial dataset retained 1,016 predictions, reflecting differences in the number of detections generated by each training approach. Confusion matrices for each condition were generated using the independent test set.

Canopy complexity scoring

Canopy structural complexity was assessed using an ordinal rubric based on drone-observable morphological traits rather than botanical complexity alone. Three qualitative parameters were evaluated: (1) vertical layering, defined as visible height variation within the canopy (from minimal in rice and wheat to tall, multi-level structures in sugarcane, 3–6 m); (2) structural heterogeneity, defined as the presence of distinct visible features such as stalks, branches, or reproductive structures; and (3) viewing-angle dependence, defined as whether canopy appearance differs substantially between nadir and oblique views. Each parameter was evaluated qualitatively and used to assign each crop to one of three ordinal canopy complexity classes: low (1), medium (2), or high (3). The resulting assignments are reported in the Results section.

Statistical analysis

Accuracy improvement was defined as the difference in accuracy between multi-perspective and aerial imaging for each crop. A monotonic relationship between canopy complexity and improvement was tested using Spearman rank correlation. Mean improvements across low, medium, and high complexity groups were compared using a one-way ANOVA test. Because canopy complexity was assigned as an ordinal class (1-3), all correlation analyses used rank-based (Spearman) statistics. Statistical analyses were performed using the size-filtered predictions in Python 3.10 using SciPy v1.11 and pandas v2.0.

Received: July 6, 2025
Accepted: December 8, 2025
Published: September 1, 2026

REFERENCES

1. Norris-Tull, D. "Defining the Problem of Invasive Species." Management of Invasive Plants in the Western USA, July 2020, www.invasiveplantswesternusa.org/. Accessed 12 Jan 2025.
2. Nashville Tree Conservation Corps. "Understanding the Dangers of Invasive Plants." Nashville Tree Conservation Corps, 13 Oct 2024, www.nashvilletreeconservationcorps.org/treenews/dangers-of-invasive-plants. Accessed 12 Jan 2025.
3. Potter, Kevin M., et al. "How Invaded Are Hawaiian Forests? Non-Native Understory Tree Dominance Signals Potential Canopy Replacement." Landscape Ecology, vol. 38, 25 April 2023, pp. 3903-3923. https://doi.org/10.1007/s10980-023-01662-6.
4. Hughes, R. Flint, et al. "Biology and Impacts of Pacific Islands Invasive Species: Falcataria falcata (Miquel) Barneby and Grimes." *Pacific Science*, vol. 78, no. 1, 7 October 2024. https://doi.org/10.2984/78.1.5.
5. Hawaiʻi Invasive Species Council. "Strategic Plan for the Control and Management of Albizia in Hawaiʻi." State of Hawaiʻi, Department of Land and Natural Resources, 2018, dlnr.hawaii.gov/hisc/files/2018/01/Strategic-Plan-for-the-Control-and-Management-of-Albizia-In-Hawaii.pdf. Accessed 12 Jan 2025.
6. Fennell, M., et al. "Japanese Knotweed (Fallopia japonica): An Analysis of Capacity to Cause Structural Damage and a Review of Control Methods." PeerJ, vol. 6, 2018, e5246. https://doi.org/10.7717/peerj.5246.
7. Williams, D. A., et al. "Emerging Insights on Brazilian Pepper Tree (Schinus terebinthifolius)." Frontiers in Plant Science, vol. 7, 2016, Article 712. https://doi.org/10.3389/fpls.2016.00712.
8. Pimentel, D., et al. "Update on the Environmental and Economic Costs Associated with Alien-Invasive Species in the United States." Ecological Economics, vol. 52, no. 3, 2005, pp. 273–288. https://doi.org/10.1016/j.ecolecon.2004.10.002.
9. Mehta, S. V., et al. "Optimal Detection and Control Strategies for Invasive Species Management." Ecological Economics, vol. 61, no. 2–3, 2007, pp. 237–245. https://doi.org/10.1016/j.ecolecon.2006.10.024.
10. van Leeuwen, M., et al. "Automated Reconstruction of Tree and Canopy Structure for Modeling the Internal Canopy Radiation Regime." Remote Sensing of Environment, vol. 136, 2013, pp. 286–300. https://doi.org/10.1016/j.rse.2013.05.021.
11. Goel, N. S. "Models of Vegetation Canopy Reflectance and Their Use in Estimation of Biophysical Parameters from Reflectance Data." Remote Sensing Reviews, vol. 4, 1988, pp. 1–212. https://doi.org/10.1080/02757258809532105.
12. Yu, K., et al. "Understanding the Impact of Vertical Canopy Position on Leaf Spectra and Functional Traits." Remote Sensing, vol. 13, 2021, Article 5057. https://doi.org/10.3390/rs13245057.
13. Zhang, C., et al. "Tree Species Classification Using Plant Functional Traits and Canopy Structure Information." Remote Sensing, vol. 14, 2022, Article 6227. https://doi.org/10.3390/rs14246227.
14. Zaka, M. M., and Samat, A. "Advances in Remote Sensing and Machine Learning Methods for Invasive Plants Study: A Comprehensive Review." Remote Sensing, vol. 16, no. 20, 2024, Article 3781. https://doi.org/10.3390/rs16203781.
15. Jensen, T., et al. "Employing Machine Learning for Detection of Invasive Species Using Sentinel-2 and AVIRIS Data: The Case of Kudzu in the United States." Sustainability, vol. 12, no. 9, 2020, Article 3544. https://doi.org/10.3390/su12093544.
16. Youssef, A., et al. "DeepForest: Sensing into Self-Occluding Volumes of Vegetation." arXiv Preprint, 2025, arXiv:2502.02171. https://arxiv.org/abs/2502.02171.
17. Reagan, J. "Making Sense of Drone-Based Sensors." dronegenuity, www.dronegenuity.com/making-sense-of-drone-based-sensors/. Accessed 12 Jan 2025.
18. Chisholm, R. A., et al. "UAV LiDAR for Below-Canopy Forest Surveys." Journal of Unmanned Vehicle Systems, vol. 1, no. 1, 2013, pp. 61-68. https://doi.org/10.1139/juvs-2013-0017.
19. "Agriculture Drones: Farming Drone for Crop Monitoring." JOUAV, 17 Dec 2024, jouav.com/blog/agriculture-drone.html. Accessed 12 Jan 2025.
20. Looby, C. "Researchers Turn to Drones for That Big-Picture View of the Forest Canopy." Mongabay, 21 Mar 2022, news.mongabay.com/2022/03/researchers-turn-to-drones-for-that-big-picture-view-of-the-forest-canopy/. Accessed 12 Jan 2025.
21. Ghanbari Parmehr, E., and Amati, M. "Individual Tree Canopy Parameters Estimation Using UAV-Based Photogrammetric and LiDAR Point Clouds in an Urban Park." Remote Sensing, vol. 13, no. 11, 2021, Article 2062. https://doi.org/10.3390/rs13112062.
22. Boesch, G. "Object Detection: The Definitive 2025 Guide." viso.ai, 4 Oct 2024, viso.ai/deep-learning/object-detection/. Accessed 12 Jan 2025.
23. "Object Detection in Agriculture." Saiwa, 16 Mar 2024, saiwa.ai/blog/object-detection-in-agriculture/. Accessed 12 Jan 2025.
24. Lin, T.-Y., et al. "Focal Loss for Dense Object Detection." arXiv Preprint, 2017, arXiv:1708.02002. https://arxiv.org/abs/1708.02002.
25. Chen, Y., et al. "IPMCNet: A Lightweight Algorithm for Invasive Plant Multiclassification." Agronomy, vol. 14, no. 2, 2024, Article 333. https://doi.org/10.3390/agronomy14020333.
26. Lu, D., and Wang, Y. "MAR-YOLOv9: A Multi-Dataset Object Detection Method for Agricultural Fields Based on YOLOv9." PLOS ONE, vol. 19, no. 10, 2024, Article e0307643. https://doi.org/10.1371/journal.pone.0307643.
27. Pashaei, M. "Applications of Deep Learning and Multi-

https://doi.org/10.59720/25-277

Perspective 2D/3D Imaging Streams for Remote Terrain Characterization of Coastal Environments." Dissertation, Texas A&M University–Corpus Christi, 2021. https://www.proquest.com/openview/c3ddf075d2a943d055512e9ae99235d8/1

28. Shaik, H. "A Study on Creating Digital Twin Foliage Representation through Computer Vision, Aerial Image Analysis and Machine Learning." Dissertation, 2024. ProQuest. https://www.proquest.com/openview/cad5069f04a6cb6d99a7c001a6ae9097/1
29. Wang, F., et al. "Three-Dimensional Reconstruction of Soybean Canopy Based on Multivision Technology for Calculation of Phenotypic Traits." Agronomy, vol. 12, no. 3, 2022, Article 692. https://doi.org/10.3390/agronomy12030692.
30. Onwudinjo, K.C. "Evaluating the Performance of Multi-Rotor UAV-SfM Imagery in Assessing Simple and Complex Forest Structures." Dissertation, University of Cape Town, 2019. https://open.uct.ac.za/items/70b2ae37-482f-4559-9bff-e7272cddb31c
31. Jaiswal, A. "Agriculture Crop Images." Kaggle, 2020, https://www.kaggle.com/datasets/aman2000jaiswal/agriculture-crop-images. Accessed 12 Jan 2025.
32. He, K., et al. "Deep Residual Learning for Image Recognition." arXiv Preprint, 2015, arxiv.org/abs/1512.03385.
33. Weinstein, B. G., et al. "Cross-Site Learning in Deep Learning RGB Tree Crown Detection." Ecological Informatics, vol. 56, 2020, Article 101061. https://doi.org/10.1016/j.ecoinf.2020.101061.
34. Weinstein, B. G., et al. "Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks." Remote Sensing, vol. 11, no. 11, 2019, Article 1309. https://doi.org/10.3390/rs11111309.
35. Weinstein, B. G., et al. "The Cropmodel." DeepForest Documentation, 1 June 2019, deepforest.readthedocs.io/en/latest/user_guide/03_cropmodels.html. Accessed 12 Jan 2025.

Copyright: © 2026 Yadati and He. All JEI articles are distributed under the Creative Commons Attribution Noncommercial No Derivatives 4.0 International License. This means that you are free to share, copy, redistribute, remix, transform, or build upon the material for any purpose, provided that you credit the original author and source, include a link to the license, indicate any changes that were made, and make no representation that JEI or the original author(s) endorse you or your use of the work. The full details of the license are available at https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en.