DENDROBIOLOGY

2026, vol. 95, 156–169

https://doi.org/10.12657/denbio.095.011

Nguyen Thanh Tuan, Phan Van Tuan, Nguyen Thanh Hung, Duong Tien Duc*

Individual tree detection and crown segmentation for *Pinus kesiya* using UAV–LiDAR and deep learning models

Received: 7 January 2026; Accepted: 14 May 2026

Abstract: Unmanned aerial vehicle (UAV) is increasingly used for individual tree crown detection and segmentation; however, performance remains limited in structurally heterogeneous stands due to crown overlap and variable crown geometry. This study evaluates and compares three advanced deep learning approaches for individual tree crown detection and segmentation from orthophoto RGB imagery in *Pinus kesiya* plantations, including YOLOv8m-Seg, YOLOv11m-Seg, and an integrated YOLOv12m–SAM2 framework. The optimal model is subsequently selected to derive individual tree attributes by integrating LiDAR data and orthophoto RGB imagery. The results revealed significant differences in performance among the evaluated models. YOLOv11m-Seg framework achieved the highest detection performance, with mAP50 = 0.62 (0.59–0.64), mAP50–95 = 0.290 (0.287–0.294), and an F1-score of 0.61 (0.58–0.63), indicating accurate localization of tree crowns. For the segmentation task, the model achieved F1-score = 0.59 (0.56–0.61), mAP50 = 0.60 (0.58–0.62), and mAP50–95 = 0.26 (0.25–0.27), outperforming YOLOv12m + SAM2 and YOLOv8m-Seg at the IoU > 0.5 threshold. Finally, the accuracy of individual tree attribute extraction from LiDAR data and RGB imagery using YOLOv11m-Seg was validated through comparison with ground measurements. The results showed a higher agreement between reference data and UAV–LiDAR data for tree height than for crown diameter with R^2 values ranging from 0.85 to 0.93 for tree height and from 0.58 to 0.86 for tree crown diameter. Overall, the findings suggest that YOLOv11m-Seg provides a reliable approach for extracting individual tree attributes from UAV–LiDAR data in *Pinus kesiya* plantations.

Keywords: drone, forest inventory, Segment Anything Model (SAM), You Only Look Once (YOLO)

Addresses: N. T. Tuan, P. Van Tuan, Faculty of Forestry, Dong Nai Campus, Vietnam National University of Forestry, Bien Hoa 810000, Vietnam, e-mail: nttuan@vnuf2.edu.vn
N. T. Hung, Ho Chi Minh University of Natural Resources and Environment, Ho Chi Minh City, 700000, Vietnam
D. T. Duc, Institute of Tropical Forest Research and Development, Ha Noi, 100000, Vietnam
* corresponding author

Introduction

Forests play a vital role in climate regulation and the global carbon cycle (Hurteau, 2021; Psistaki et al., 2024). In recent decades, natural forests have been severely degraded by war, natural disasters, and land-use conversion, leading to a significant decline in their extent and ecological functions (Dang, 2022; Hallaj et al., 2024; Ullah, 2025). In this context, plantation forests not only provide timber resources but also contribute substantially to carbon sequestration and climate regulation (Di Sacco et al., 2021). Well-managed plantation forests can partially compensate for the loss of natural forests by enhancing carbon storage and supporting climate change mitigation efforts (Diao et al., 2022; Wang et al., 2025). In plantation forest management, the assessment of economic and environmental benefits, such as timber yield and carbon sequestration capacity, is an essential tasks. Accurate and efficient methods for estimating wood production and carbon stocks are therefore critical for improving plantation forest management and decision-making (Baskent et al., 2025; Hua et al., 2025; Suarez-Fernandez et al., 2025). Indeed, the carbon stored in a tree can be estimated using allometric equations that relate biomass to structural attributes such as species, tree height, and diameter at breast height (Jucker et al., 2022). However, conventional field-based measurements are labor-intensive, time-consuming, and often limited by terrain accessibility (Jurjević et al., 2020; Kim et al., 2024; Xin et al., 2025). Accurately measuring tree height is particularly challenging in closed-canopy forests, as dense and overlapping crowns often obscure treetop; in many cases, only lateral branches are visible from the ground, leading to systematic underestimation (Laurin et al., 2019; Saliu et al., 2021).

In recent years, unmanned aerial vehicle (UAV) platforms equipped with LiDAR sensors and high-resolution optical cameras have emerged as powerful tools for acquiring detailed three-dimensional information on forest structure (Bumbálek et al., 2025; Fu et al., 2024; Marcello et al., 2024; Wang et al., 2025). Individual tree identification is a fundamental component of modern forest inventories, enabling more accurate estimation of stem volume, above-ground biomass, and carbon stocks, while also supporting biodiversity conservation, forest health monitoring, and sustainable forest management (Satama-Bermeo et al., 2025; Sengun et al., 2025a). UAV–LiDAR enables precise characterization of canopy height and vertical structure, whereas optical imagery provides fine spatial detail for delineating individual tree crowns. However, the reliable extraction of individual tree attributes from these datasets remains challenging, particularly in dense, heterogeneous-canopy forests (Dersch et al., 2023; Nikitina, 2024; Rahman et al., 2025; Zhao et al., 2023).

To address these challenges, many countries have increasingly adopted machine learning (ML) to develop automated systems, robots, and decision-support tools that assist humans in forest resource monitoring and management (Abbas & Damaševičius, 2025). In particular, deep learning (DL) has demonstrated outstanding performance in object detection and instance segmentation tasks, and is being increasingly applied in forestry-related research and operational applications (He et al., 2025; Mendes et al., 2023; Reisi Gahrouei et al., 2024; Wołk & Tatara, 2024). Models such as You Only Look Once (YOLO) enable rapid detection of individual trees. In contrast, the Segment Anything Model (SAM) has demonstrated strong generalization capabilities, enabling flexible and data-efficient segmentation (Ali & Zhang, 2024; Dong et al., 2024; Pandey et al., 2023; Que et al., 2026; Ramos & Sappa, 2025). Combining object detection and segmentation models (e.g., YOLO + SAM) offers new opportunities to improve the accuracy of tree crown extraction (Cabral et al., 2025; Teng et al., 2025).

Pinus kesiya, a dominant plantation species in cooler and drier montane regions, particularly in the Central Highlands of Vietnam, plays a crucial role in timber production and the provision of ecosystem services. Despite its importance, there has been limited research on the application of deep learning and UAV–LiDAR data for automated extraction of tree attributes, aimed at improving the efficiency of timber volume estimation, carbon stock assessment, and structural characterization of this species in Vietnam. Therefore, this study aims to (1) compare the accuracy, stability, and reliability of three deep learning architectures (YOLOv8m, YOLOv11m, and YOLOv12m-SAM2) for tree detection and crown segmentation; (2) assess the accuracy of tree attribute extraction derived from UAV–LiDAR data by comparing them with field measurements across different stand conditions. The proposed framework provides a cost-effective, scalable solution for semi-automatic forest inventory and advances the application of deep learning and UAV–LiDAR for extracting stand characteristics, thereby supporting carbon estimation for *Pinus kesiya* plantations in Vietnam.

Material and methods

Study site and data collection

The study area is located within *Pinus kesiya* plantation forest in La Ba Co., Ltd., Don Duong District, Lam Dong province. The study area extends from 11.72°–11.74°N and 108.46°–108.49°E (Fig. 1). In

this study, UAV–LiDAR data were collected on November 25, 2023, using the DJI Matrice 300 RTK with a Zenmuse L1 camera, at a flight altitude of approximately 100 m in autonomous mode. The camera was oriented in a nadir (vertical) position, with forward and side overlaps both set to 80% to ensure sufficient image coverage for orthophoto generation. The Zenmuse L1 is an all-in-one aerial mapping sensor that integrates a LiDAR module, an RGB camera, and a high-precision IMU. The sensor supports up to three returns, enabling double and triple echoes that can increase the point density to as much as 240,000 points per second. LiDAR module has a maximum detection range of 450 m under 80% reflectivity at 0 klx, and 190 m under 10% reflectivity at 100 klx.

During UAV–LiDAR data preprocessing, the LiDAR point cloud (~236 pts/m^2) was first classified into ground and non-ground returns using a progressive TIN densification filtering algorithm implemented in LiDAR 360 (version 5.4). The Triangulated Irregular Network (TIN) interpolation method was subsequently applied to the classified points to generate the Digital Elevation Model (DEM) and Digital Surface Model (DSM). The grid resolution was defined by the XSize and YSize parameters, both set to 0.05 m, representing the sampling interval of the output raster. Finally, a Canopy Height Model (CHM) was derived by subtracting the DEM from the DSM using raster differencing.

The photo datasets were processed in Agisoft Metashape (version 2.2.3) following a standard photogrammetric workflow. Images were aligned using high accuracy, with generic and reference pair preselection enabled. Dense point clouds were generated using high quality and mild depth filtering. A Digital Surface Model (DSM) was then created from the dense cloud, and the orthomosaic was generated using mosaic blending mode with surface-based projection. All outputs were referenced to the projected coordinate system EPSG:5899. The final orthophotos had a spatial resolution of approximately 0.03 m per pixel. The orthomosaic of the study area was then divided into 598 separate image tiles, each with a resolution of 716 × 716 pixels. From these, 170 tiles containing the occurrence of *Pinus kesiya* were selected for tree-crown delineation. Of these, 167 tiles were used to develop the tree-crown segmentation models, while 3 tiles corresponding to three 500 m^2

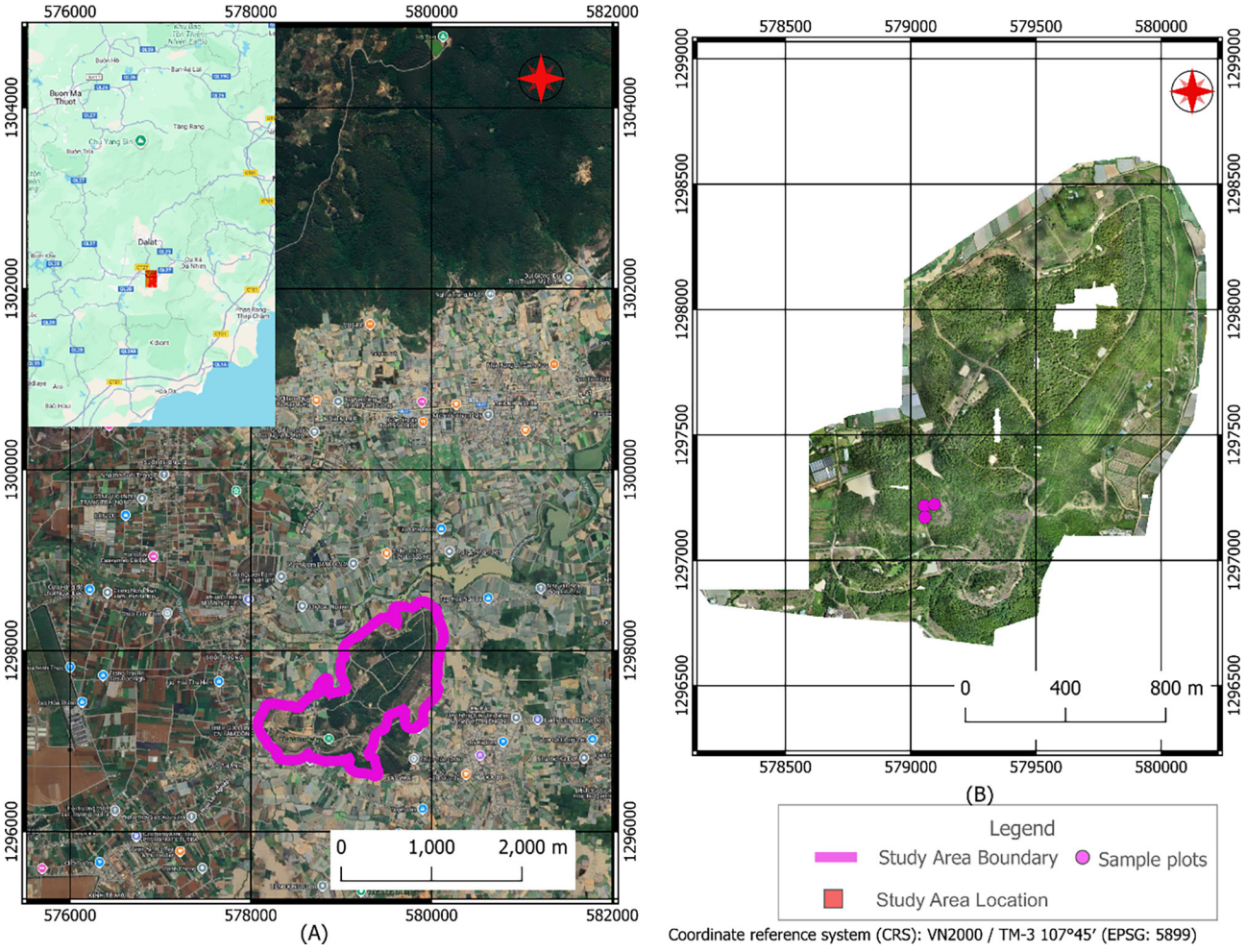

Fig. 1. Study area and datasets (A: The location of the study site; B: UAV RGB images)

sample plots (22.4 m × 22.4 m) were used to test the accuracy of the deep learning–based tree-crown segmentation. Three sample plots were selected within an experimental thinning site to represent different stand densities and spatial tree distributions resulting from thinning treatments. This design enables evaluation of model performance under varying forest structural conditions.

After UAV image acquisition, we conducted 3 sampled plots of 500 m^2 (22.4 m × 22.4 m) to collect forest stand characteristics, including the position, height and crown width of each tree in the sampled plot. The geographic positions of individual trees were measured using a Geomate SG20AR RTK GNSS system. Dense forest canopy conditions made it difficult to receive sufficient GNSS signals for accurate positioning. However, in this study, the measurements achieved an accuracy of better than 20 cm. Tree height was measured using different approaches depending on crown visibility. When the tree apex was clearly visible, total tree height was measured directly by LaserTech Criterion RD 1000. Conversely, the apex was obscured by overlapping canopies, field operators used a laser rangefinder from a nearby position, aiming the laser beam at upper branches or foliage to estimate tree height. The crown lengths of each tree in the north–south (N–S) and east–west (E–W) directions were measured with a tape along the trunk in both perpendicular directions. Based on these measurements and field observations, the crown shape was sketched on lined paper following the guidelines of (Davis & Richards, 1933). The sketched crown shapes were then digitized into polygon shapefiles (SHP) in Qgis 3.44 for use as reference data in model evaluation.

Instance-segmentation annotations were manually delineated for each individual tree crown across 167 image tiles using Roboflow. Crown boundaries were defined based on visual interpretation of UAV RGB imagery, following a consistent annotation protocol. To ensure annotation quality, all annotators were trained using the same guidelines. In cases where tree crowns were clearly separated, boundaries were delineated along visible crown edges. For overlapping canopies, crown delineation was guided by crown apex position, texture, and color differences in the RGB imagery, and further supported by LiDAR-derived data, including the canopy height model (CHM) and point cloud structure, to distinguish adjacent trees. No automated or AI-assisted annotation methods were applied. In total, 1623 individual tree crowns were annotated.

Methods used

In the first step, we employed multiple deep learning architectures for individual tree detection and crown segmentation from UAV imagery, including YOLOv8, YOLOv11, and a hybrid YOLO + SAM approach. In this study, we employed the medium-scale (m) variants of YOLOv8, YOLOv11, and YOLOv12. The model was trained for 150 epochs with a batch size of 16 and an input image resolution of 640 × 640 pixels. Training was performed using stochastic gradient descent optimization with an initial learning rate of 0.01, momentum of 0.937, and weight decay of 0.0005. A cosine learning rate schedule was not applied, and early stopping was implemented with a patience of 100 epochs. The loss function followed the default YOLO segmentation formulation, combining bounding box regression, classification, and distribution focal loss components. Automatic mixed precision (AMP) was enabled to improve computational efficiency. The "m" configuration represents a balanced model size in the YOLO family, providing a compromise between computational efficiency and detection accuracy (Bumbálek et al., 2025; Jegham et al., 2024). In addtion, the m variants were particularly appropriate for our dataset, which contains small and densely distributed targets, where higher feature capacity is beneficial without incurring excessive inference time. To ensure a robust and unbiased evaluation, a five-fold cross-validation scheme was implemented (Darma et al., 2022; Pranata et al., 2023; Prousalidis et al., 2024). The dataset was randomly partitioned into five equal subsets, where in each iteration, four subsets were used for training and the remaining subset for validation. The model was trained using Google Colab T4 GPU with high RAM and Ultralytics 8.3.246, Python-3.12.12, torch-2.9.0+cu126 CUDA:0 (Tesla T4, 15095MiB). The full training configuration file is provided in the Supplementary Material (Fig. S2).

Two methodological strategies were implemented for tree-crown detection and segmentation in YOLO.

Method 1: Used end-to-end instance segmentation models, including YOLOv8 and YOLOv11. In this approach, the models directly performed both object detection and tree-crown segmentation within a single framework. Bounding boxes and crown masks were generated simultaneously during inference, allowing the models to learn crown-shape information during training. This pipeline represents a conventional one-stage/ two-stage deep learning solution for tree-crown delineation.

Method 2: YOLOv12 currently does not support instance segmentation. This method employed a two-stage workflow that integrates object detection with prompted segmentation. First, the YOLOv12-m model was trained to detect individual trees and generate bounding-box predictions. These bounding boxes were subsequently used as prompts for the SAM-2 segmentation model, which refined the crown boundaries and produced

instance-level crown masks. This approach separates the detection and segmentation processes, enabling the YOLO models to provide accurate tree localization, while SAM-2 improves crown delineation in areas where adjacent tree crowns exhibit horizontal overlap within the visible upper canopy.

All results are compared with other methods according to precision, recall, F1 score, and mean average precision (mAP). The results are also compared with each other and predict the best. The evaluation metrics used in this study are defined as follows:

Precision measures the proportion of correctly predicted trees among all predicted bounding boxes.

$$Precision = \frac{TP}{(TP + FP)} \times 100$$

Recall quantifies the proportion of true trees that were successfully detected.

$$Recall = \frac{TP}{(TP + FN)} \times 100$$

The F1-score evaluates overall detection performance by combining Precision and Recall into their harmonic mean.

$$F1 - score = \frac{2\ Precision \times Recall}{(Precision + Recall)} \times 100$$

To assess the spatial accuracy of the tree crown polygons, intersection-over-union (IoU) was calculated for the reference and target crowns. IoU is used to assess how well the predicted bounding box overlaps with the ground truth bounding box. If the result is more than 0.5 it is counted as a correct detection. Mean Average Precision (mAP) is calculated by averaging AP scores across multiple Intersection over Union (IoU) thresholds. In this study, we report two variants:

- mAP50, which is computed using a single IoU threshold of 0.5.
- mAP50–95, which is obtained by averaging the AP scores across IoU thresholds from 0.50 to 0.95, using increments of 0.05.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$$

where TP represents correctly detected trees (detections with IoU greater than 0.50), FP refers to incorrectly detected trees or false detections (IoU below 0.50 or no corresponding ground-truth tree), and FN denotes trees that were present in the reference data but not detected by the model.

Performance metrics, including mAP, Precision, Recall, and F1-score, were computed for each fold. The resulting metrics were subsequently subjected to bootstrap resampling (n = 999) to estimate the mean performance and corresponding 95% confidence intervals (CI_low and CI_high). This framework enables robust performance estimation while explicitly accounting for variability and statistical uncertainty, thereby enhancing the reliability and generalizability of the results.

In the second stage, the optimal YOLO model was selected to extract stand structural attributes from the LiDAR data within the study plots. Tree crowns were delineated via segmentation, while tree heights were derived from the Canopy Height Model (CHM), specifically extracted from the center of each delineated crown.

To validate the accuracy of tree attributes (tree height, crown diameter) derived from UAV–LiDAR data, these estimates were compared with field

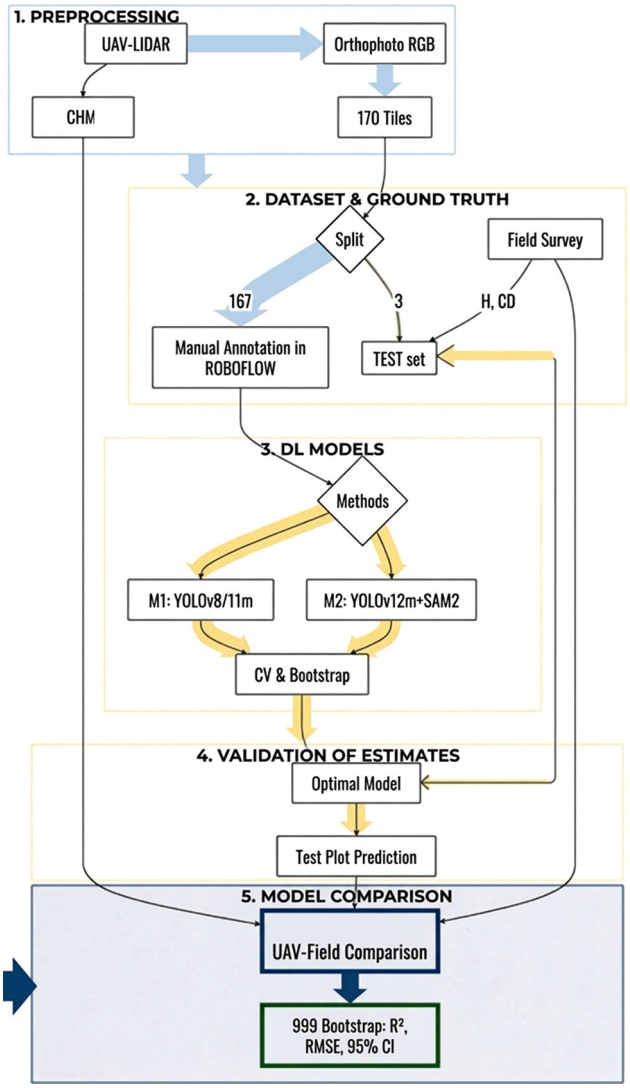

Fig. 2. Flowchart for individual tree attribute from UAV–LiDAR data using YOLO models

measurements across three plots. A regression model was developed using 999 bootstrap simulations. Model performance was evaluated using the coefficient of determination (R^2) and root mean square error (RMSE). In addition, 95% confidence intervals (CI) based on the 999 bootstrap simulations were calculated to compare the mean stand attributes derived from UAV–LiDAR data with those obtained from field measurements. All statistical analyses were conducted using R version 3.4. Figure 2 illustrates the workflow of the proposed method for extracting individual tree attributes from UAV–LiDAR data using YOLO models.

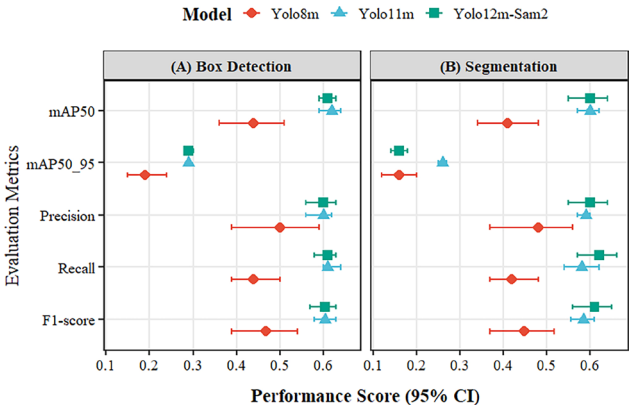

Fig. 3. Results of object detection and instance segmentation using YOLO on the validation dataset

Results

Individual tree detection and crown segmentation model training

The performance of the three deep learning architectures (YOLOv8m, YOLOv11m, and YOLOv12m-Sam2) was evaluated using 5-fold cross-validation, followed by 999 bootstrap resamples to estimate the mean and 95% confidence intervals (CI) for each metric. The results, summarized in Table S1 and illustrated in Figure 3, indicate that both YOLOv11m and YOLOv12m-Sam2 substantially outperformed the baseline YOLOv8m across all metrics in both detection and segmentation tasks.

In the bounding box detection task, the advanced models demonstrated high precision and recall stability (Fig. 2A). YOLOv11m achieved the highest mean mAP50 of 0.62 (95% CI: 0.59–0.64), followed by YOLOv12m-Sam2 with a mean of 0.61 (95% CI: 0.59–0.63). Both YOLOv11m and YOLOv12m-Sam2 reached a mean mAP50–95 of 0.29, a substantial improvement over the 0.19 recorded for YOLOv8m. The

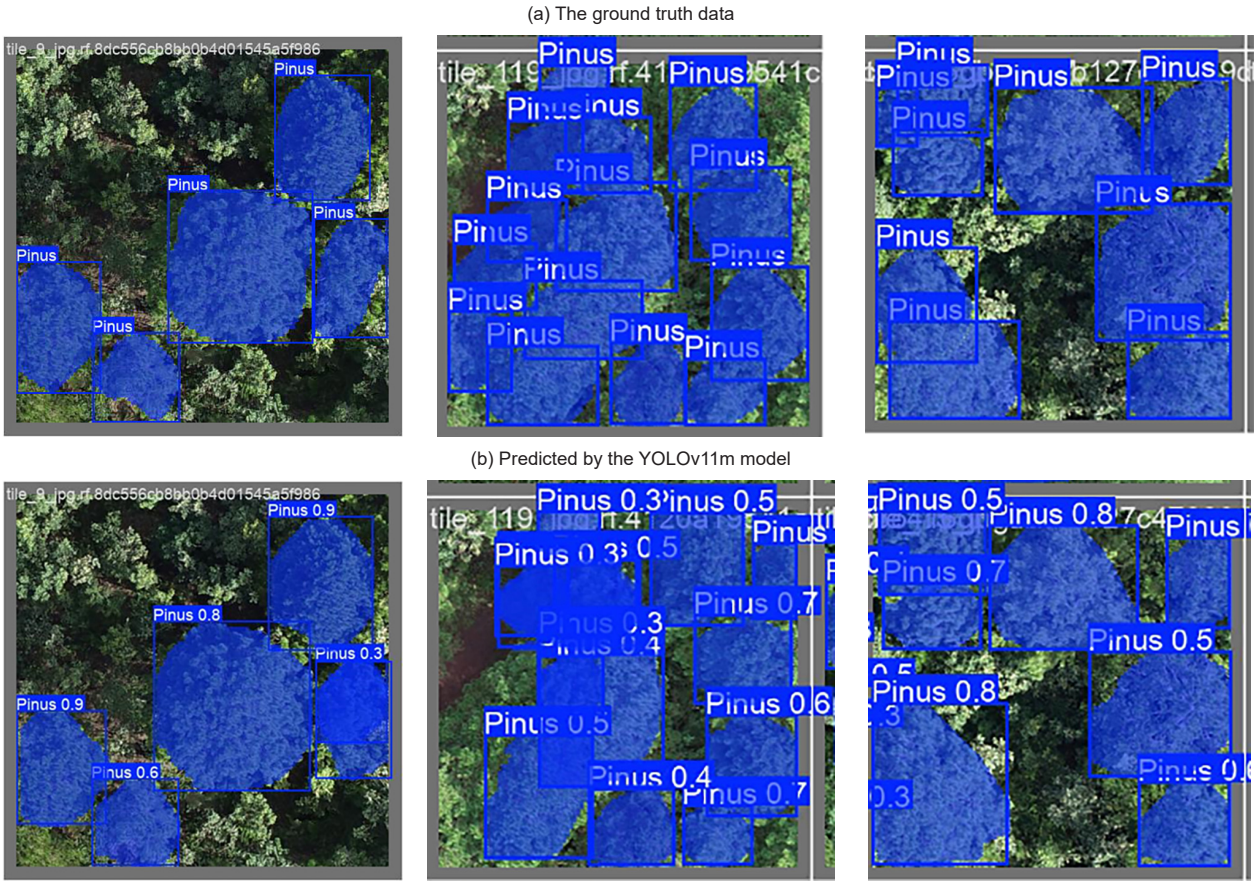

Fig. 4. Qualitative comparison on a subset of the validation dataset

F1-score for both YOLOv12m-Sam2 and YOLOv11m models was recorded at 0.61, reflecting a balanced trade-off between precision and recall, compared to 0.47 for the YOLOv8m model.

The segmentation task revealed critical differences in boundary refinement and statistical reliability among the models (Fig 2B). Both YOLOv11m and YOLOv12m-Sam2 achieved a mean mAP50 of 0.60. However, YOLOv11m demonstrated superior statistical stability with a narrower 95% CI (0.57–0.62) compared to the wider interval of YOLOv12m-Sam2 (0.55–0.64). A significant performance gap was observed in the mAP50–95 metric. YOLOv11m yielded a mean of 0.26 (95% CI: 0.25–0.27), while YOLOv12m-Sam2 and YOLOv8m only achieved 0.16. The F1-score for YOLOv11m was 0.59 (95% CI: 0.56–0.61), whereas YOLOv12m-Sam2 exhibited a mean of 0.61 but with a larger variance (95% CI: 0.56–0.65).

Notably, Yolo11m consistently exhibited narrower confidence intervals across most metrics, particularly in segmentation, suggesting greater stability and reliability compared to Yolo12m-Sam2.

The performance of tree detection and instance segmentation on the validation set using YOLOv11m model is presented in Figure 4.

Evaluation of tree attribute extraction accuracy from UAV–LiDAR data on test datasets

YOLOv11m was used to extract tree attributes from UAV–LiDAR data across three test plots representing different stand densities. Across the three plots, UAV–LiDAR extracted showed generally consistent structural patterns compared with field measurements, though with systematic deviations. Tree height estimates from UAV–LiDAR exhibited a systematic pattern of being equal to or higher than field values in plot 1 and plot 3 (9.07 m vs. 7.62 m; 10.43 m vs. 9.96 m), while showing a slight underestimation in plot 2 (10.72 m vs. 11.55 m) (Fig. 5). The 95% confidence intervals (CI) of the mean values derived from 999 bootstrap resamples indicate no significant difference in mean tree height between field-measured data and UAV–LiDAR estimates in plot 2 and plot 3. In contrast, a significant difference was observed in plot 1.

Tree crown diameters derived from UAV–LiDAR were slightly larger than field measurements in plot 1 (3.31 m vs. 2.71 m), but smaller in plot 2 and plot 3 (3.47 m vs. 4.40 m, and 3.30 m vs. 3.57 m,

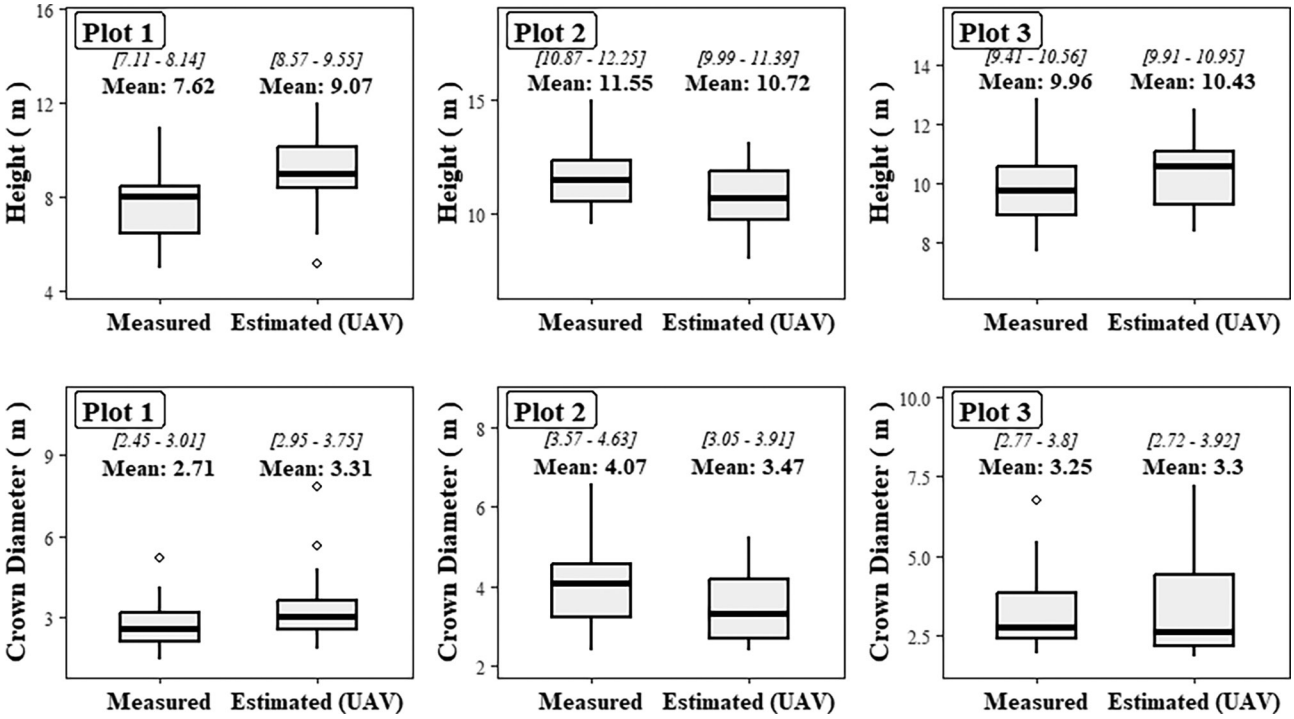

Fig. 5. Box and whisker plots of the field-measured and UAV–LiDAR estimated. The intervals represent the 95% confidence intervals (CI) of the mean values derived from 999 bootstrap resamples

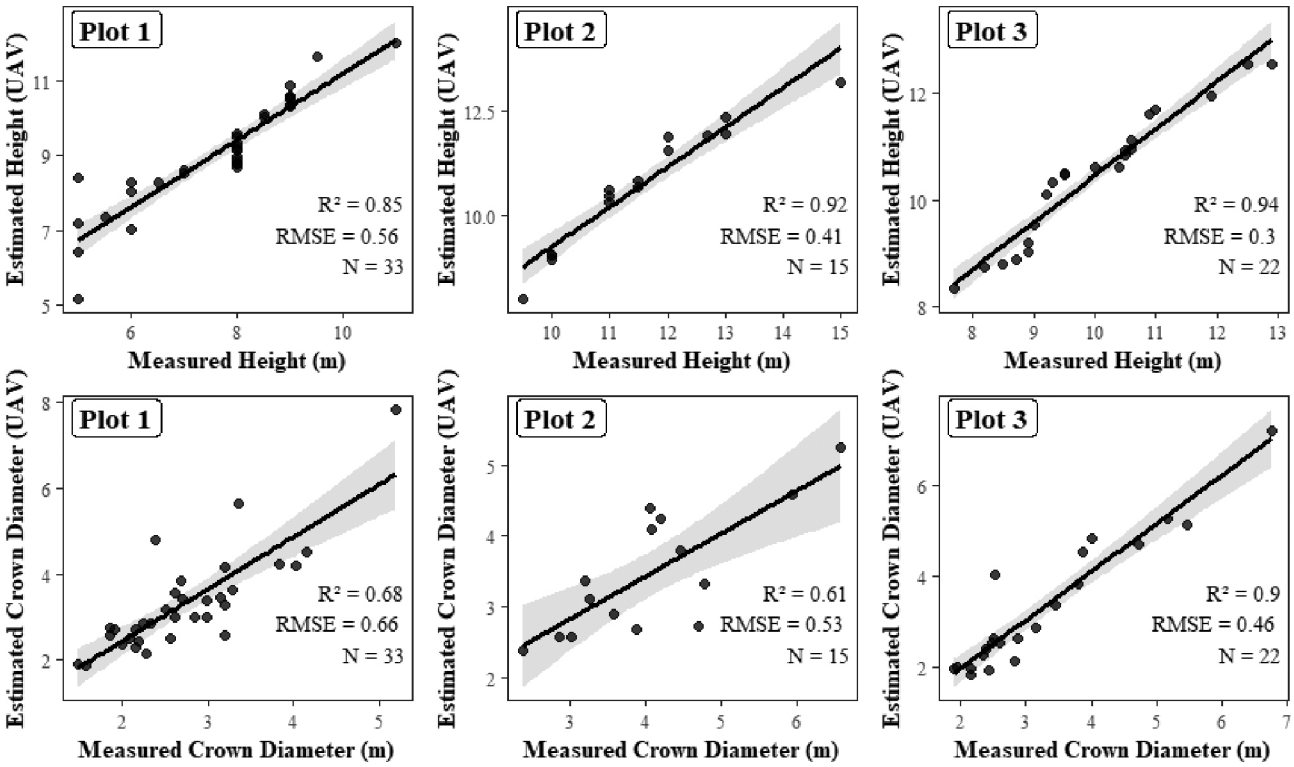

Fig. 6. Scatter plots showing the relationship between predicted individual tree attributes from the YOLOv11 model and field-measured data in *Pinus kesiya* plantation forests. The gray intervals represent the 95% confidence intervals (CI) derived from 999 bootstrap resamples

respectively). In addition, the 95% confidence intervals (CI) indicate no significant difference in mean tree crown size between field-measured data and UAV–LiDAR estimates across the three plots.

Figure 6 shows the relationship between UAV–LiDAR-estimated tree attributes and reference tree measurements. Linear regression revealed a strong relationship between estimated and measured tree heights in plot 1 ($R^2 = 0.85$), plot 2 ($R^2 = 0.93$), and plot 3 ($R^2 = 0.93$). Similarly, the estimated crown diameters showed good agreement with field measurements, with R^2 values of 0.67, 0.58, and 0.88 for plots 1, 2 and 3, respectively. In addition, RMSE values decreased from plot 1 to plot 3.

Discussion

The ability of YOLO models to utilize individual tree detection and segmentation

The experimental results reveal a clear progression in model performance from the YOLOv8 architecture to more recent iterations for individual tree crown delineation using UAV imagery. Notably, YOLOv11m demonstrates substantial improvement over YOLOv8m while maintaining performance comparable to the more complex YOLOv12–SAM2 framework. YOLOv11 is the latest evolution in the Ultralytics YOLO series, introducing architectural improvements over YOLOv8 such as replacing the C2f module with C3K2 and adding a C2PSA block to enhance feature representation. It also refines the detection head and optimizes network scaling, resulting in improved accuracy and inference efficiency (Liu et al., 2025). A key finding of this study is the superior performance of YOLOv11m under the most stringent evaluation metric. Although the models exhibit overlapping 95% confidence intervals for standard metrics such as Precision and F1-score, a distinct performance gap is observed in mAP50–95 for the segmentation task. This suggests that, while the integration of SAM2 may enhance general mask detection, YOLOv11m achieves higher spatial fidelity and more consistent intersection-over-union (IoU) performance across multiple thresholds. The non-overlapping confidence intervals for mAP50–95 further confirm the statistical significance of this difference, highlighting YOLOv11m as the most reliable model for applications requiring precise pixel-level delineation.

These findings are consistent with previous studies demonstrating the strong potential of YOLOv11 for tree detection and crown delineation in aerial imagery of Indonesia's forestry landscapes (Rato et al., 2025). Similarly, the model has been successfully applied in urban environments for accurate tree

species identification and biomass estimation (Huo et al., 2026). The YOLO-based approach effectively overcomes several limitations associated with conventional segmentation frameworks such as U-Net, DeepLabV3+, Mask R-CNN, SwinUNet, RepViT-SAM, and EdgeSAM and the Segment Anything Model (SAM). By unifying object detection and segmentation into a single architecture, the model achieves a favorable balance between accuracy and computational efficiency, making it particularly well-suited for large-scale and time-sensitive applications (Albahli, 2025; Prousalidis et al., 2024). More broadly, YOLO-based object detection frameworks have been widely recognized as efficient, scalable, and accurate solutions for individual tree detection using both UAV–RGB imagery and LiDAR data, thereby supporting improved forest inventory, monitoring, and ecosystem management (Farhan et al., 2025; Luo et al., 2024; Sengun et al., 2025b; Zhong et al., 2024).

Despite the overall performance improvements, segmentation accuracy across IoU thresholds remained relatively low on the validation dataset, with mAP50–95 values below 0.30 and an F1-score of 0.59 (Fig. 3). In contrast, higher performance was observed in the independent test plots, where mAP50–95 ranged from 0.32 to 0.58 and F1-scores from 0.70 to 0.85 (Table S.2). This discrepancy can be primarily attributed to differences in reference data quality. The training and validation datasets were generated through manual annotation in Roboflow, where tree crowns were delineated by redrawing canopy boundaries from UAV–RGB imagery. Such an approach is inherently subjective and prone to errors, particularly in areas with overlapping canopies, leading to inconsistencies in boundary definition and reduced agreement with model predictions. In contrast, the test datasets were derived from field-based surveys, where canopy delineation was supported by in situ observations, resulting in more accurate and consistent crown boundaries. Similar findings have been reported in recent studies, which demonstrate that annotation quality significantly influences the performance of supervised machine learning models, particularly in object detection and instance segmentation tasks (Agnew et al., 2024; Alhazmi et al., 2021; Ma et al., 2022). These results underscore the sensitivity of segmentation performance to annotation quality and highlight the importance of high-precision reference data for robust model evaluation.

The performance of individual tree segmentation in this study is generally consistent with results reported in previous research on Pinus stands. For example Nikitina (2024) reported that the spread of model quality values in pine forests in Russia ranged from 0.53 to 0.96, depending on stand structure and crown overlap conditions. In comparison, the F1-scores obtained in our study were 0.59 for the validation dataset and 0.70 to 0.85 for the test datasets, which fall within this reported performance range, indicating comparable model reliability. In contrast, substantially lower detection accuracies have been reported in more structurally complex forests, including Australian mixed-species stands (Bunting & Lucas, 2006), mixed-conifer forests (Jakubowski et al., 2013), and Scots pine forests (Kaartinen et al., 2012), where accuracies often remain below 30%, as well as in natural mangroves about 46% (Yin & Wang, 2019). These comparisons highlight the strong potential of integrating YOLO-based models with UAV–LiDAR data for accurate individual tree crown delineation in *Pinus kesiya* plantations.

Effect of stand density on tree attribute extraction performance from UAV–LiDAR data

Errors in crown characterization and stand density estimation are widely reported in studies of individual tree crown detection and delineation from UAV data (Dersch et al., 2023; Zhao et al., 2023). In plantation forests, errors in crown attributes often occur because the training data do not fully capture variation, especially for smaller crowns and trees growing under different site conditions (Chadwick et al., 2020b; Zheng et al., 2021). Dersch et al. (2023) also reported difficulties in distinguishing small trees (crown size < 15 m^2) using UAV-based multispectral imagery and LiDAR data in coniferous forests. Stand density errors are also commonly reported, including crown overlap (Gomez Selvaraj et al., 2020; Mo et al., 2021) and shadow-induced misclassification (Zheng et al., 2021). Previous studies have also shown that detection accuracy generally declines as forest structural complexity and stand density increase (Yin & Wang, 2019). Consistent with these findings, our results indicate reduced model performance under dense conditions (plot 1) and in highly clustered stands (plot 2), whereas higher accuracy was achieved in more open forest conditions (plot 3), as illustrated in Figure S1.

In our study, the F1-score on the test data ranged from 0.71 to 0.77 across plots 1 and 2, whereas the $mAP_{50–95}$ remained below 0.40. This outcome suggests that, while the model is effective in locating tree crowns, it encounters difficulties in precisely delineating crown boundaries at higher overlap thresholds. The reduction in segmentation performance is likely associated with increased canopy complexity in dense stands, including crown overlap among neighboring trees and irregular crown shapes, which can lead to boundary fragmentation and partial crown delineation (Chadwick et al., 2020a; Hao et al., 2021; Lou et al., 2022; Selvaraj et al., 2020;

Zhao et al., 2023; Zheng et al., 2021). The comparison between field measurements and UAV–LiDAR extractions suggests a potential influence of stand density on the performance of the extraction workflow and the YOLOv11-based detection and segmentation model. The plot with the highest tree density (36 trees per plot, equivalent to approximately 720 trees/ha) shows the largest discrepancies in crown size (RMSE = 0.66 m), and tree height (0.56 m). This result suggests that high canopy overlap and crown interference in dense stands reduce detection accuracy and increase structural estimation bias, particularly for smaller or suppressed trees. In contrast, medium-density plots (15–25 trees per plot, below 500 trees/ha) show much closer agreement between field and UAV–LiDAR measurements (RMSE = 0.30–0.41 for tree height and 0.46–0.53 for crown diameter). This indicates that the extraction process is more reliable in stands with less crown competition and better tree separation. The influence of stand density on detection and segmentation performance is also consistent with results reported in previous studies. Windrim and Bryson (2020) evaluated Faster R–CNN in two pine stands with stem densities of 400 and 600 stems/ha and observed a decline in performance from an F1-score of 93% to 76% as stand density increased. This agreement reinforces the notion that lower canopy overlap facilitates clearer crown delineation and more reliable tree-level metrics. These results indicate that crown segmentation in complex canopy structures remains challenging. Future improvements may include refining post-processing approaches and integrating structural information, such as canopy height models in the training model, to better separate adjacent crowns and improve segmentation accuracy in dense forest conditions.

This study is limited to a single species and a single study area and relies on a relatively small dataset, with only three sample plots used for testing due to constraints in field data acquisition. Future work should include a larger number of plots distributed across diverse stand conditions, such as different stand densities, ages, and site qualities, to further evaluate the robustness and transferability of the proposed approach. Although the results provide useful insights into individual tree detection and crown segmentation for *Pinus kesiya* plantations in Lam Dong province, Vietnam, their applicability to other forest types, regions, or species remains uncertain and requires further validation.

Conclusion

This study demonstrates that advanced deep learning models substantially improve individual tree detection and crown segmentation from UAV data, with YOLOv11m providing the most consistent and reliable performance. UAV–LiDAR derived tree height and crown diameter showed strong agreement with field measurements. In addition, stand density appears to be an important factor influencing extraction accuracy, as higher-density conditions may increase crown overlap, thereby reducing detection completeness and segmentation precision.

The proposed workflow improves segmentation reliability under limited training data conditions, highlighting the adaptability of deep learning models in plantation forest environments. These findings demonstrate the potential of a one-stage deep learning object detection framework to enhance operational tree crown mapping and forest structural characterization compared with traditional field-based inventory methods, particularly in managed forest ecosystems.

Author Contributions

Nguyen Thanh Tuan: Conceptualization, Investigation, Data curation, Methodology, Formal analysis, Writing original draft and manuscript revision; Phan Van Tuan: Conceptualization, Investigation. Nguyen Thanh Hung and Duong Tien Duc: manuscript revision. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Institutional review board statement
Not applicable.

Informed consent statement
Not applicable.

Data availability statement
Not applicable.

Conflicts of interest
The authors declare no conflict of interest.

References

Abbas I & Damaševičius R (2025) Tree detection in RGB satellite imagery using YOLO-based deep learning models. Computers, Materials & Continua 85: 483–502. doi:10.32604/cmc.2025.066578.

Agnew C, Scanlan A, Denny P, Grua EM, Ven Pvd & Eising C (2024) Annotation quality versus quantity for object detection and instance segmentation. IEEE Access 12: 140958–140977. doi:10.1109/ACCESS.2024.3467008.

Albahli S (2025) A robust YOLOv8-based framework for real-time melanoma detection and segmentation with multi-dataset training. Diagnostics 15: 691–713. doi:10.3390/diagnostics15060691.

Alhazmi K, Alsumari W, Seppo I, Podkuiko L & Simon M (2021) Effects of annotation quality on model performance: 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 63–67. doi:10.1109/ICAIIC51459.2021.9415271.

Ali ML & Zhang Z (2024) The YOLO framework: a comprehensive review of evolution, applications, and benchmarks in object detection. Computers 13: 336–373. doi:10.3390/computers13120336.

Baskent EZ, Kašpar J & Baskent H (2025) Implications of carbon management with forest plantation on understocked, degraded and bare forests: Simulated long-term dynamics between timber production and carbon sequestration. Renewable Energy 242: 122437. doi:10.1016/j.renene.2025.122437.

Bumbálek R, Ufitikirezi JdDM, Umurungi SN, Zoubek T, Kuneš R, Stehlík R & Bartoš P (2025) Computer vision in precision livestock farming: benchmarking YOLOv9, YOLOv10, YOLOv11, and YOLOv12 for individual cattle identification. Smart Agricultural Technology 12: 101208. doi:10.1016/j.atech.2025.101208.

Bunting P & Lucas R (2006) The delineation of tree crowns in Australian mixed species forests using hyperspectral Compact Airborne Spectrographic Imager (CASI) data. Remote Sensing of Environment 101: 230–248. doi:10.1016/j.rse.2005.12.015.

Cabral R, Santos R, Correia JAFO & Ribeiro D (2025) A Hybrid YOLO and segment anything model pipeline for multi-damage segmentation in UAV inspection imagery. Sensors 25: 6568. doi:10.3390/s25216568.

Chadwick AJ, Goodbody TR, Coops NC, Hervieux A, Bater CW, Martens LA, White B & Röeser D (2020) Automatic delineation and height measurement of regenerating conifer crowns under leaf-off conditions using uav imagery. Remote Sensing 12: 4104–4130. doi:10.3390/rs12244104.

Dang TKP (2022) The discourse of forest cover in Vietnam and its policy implications. Sustainability 14: 10976. doi:10.3390/su141710976.

Davis T & Richards P (1933) The vegetation of Moraballi Creek, British Guiana: an ecological study of a limited area of tropical rain forest. Part I. The Journal of Ecology: 350–384. doi:10.2307/2256587

Dersch S, Schoettl A, Krzystek P & Heurich M (2023) Towards complete tree crown delineation by instance segmentation with Mask R–CNN and DETR using UAV-based multispectral imagery and lidar data. ISPRS Open Journal of Photogrammetry and Remote Sensing 8: 100037. doi:10.1016/j.ophoto.2023.100037.

Di Sacco A, Hardwick KA, Blakesley D, Brancalion PH, Breman E, Cecilio Rebola L, Chomba S, Dixon K, Elliott S, Ruyonga G, Smith P, Smith RJ & Antonelli A (2021) Ten golden rules for reforestation to optimize carbon sequestration, biodiversity recovery and livelihood benefits. Global Change Biology 27: 1328–1348. doi:10.1111/gcb.15498.

Diao J, Liu J, Zhu Z, Wei X & Li M (2022) Active forest management accelerates carbon storage in plantation forests in Lishui, southern China. Forest Ecosystems 9: 100004. doi:10.1016/j.fecs.2022.100004.

Dong G, Wang Z, Chen Y, Sun Y, Song H, Liu L & Cui H (2024) An efficient segment anything model for the segmentation of medical images. Scientific reports 14: 19425. doi:10.1038/s41598-024-70288-8.

Darma IWAS, Suciati N & Siahaan D (2022) GFF-CARVING: Graph feature fusion for the recognition of highly varying and complex balinese carving motifs. IEEE Access 10: 129217–129230. doi:10.1109/ACCESS.2022.3228382.

Farhan M, Akhtar MN & Bakar EA (2025) Efficient real-time palm oil tree detection and counting using YOLOv8 deployed on edge devices. Journal of Umm Al-Qura University for Engineering and Architecture 16: 1293–1308. doi:10.1007/s43995-025-00164-7.

Fu H, Zhao H, Jiang J, Zhang Y, Liu G, Xiao W, Du S, Guo W & Liu X (2024) Automatic detection tree crown and height using Mask R-CNN based on unmanned aerial vehicles images for biomass mapping. Forest Ecology and Management 555: 121712. doi:10.1016/j.foreco.2024.121712.

Gomez Selvaraj M, Vergara A, Montenegro F, Alonso Ruiz H, Safari N, Raymaekers D, Ocimati W, Ntamwira J, Tits L, Omondi AB & Blomme G (2020) Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin. ISPRS Journal of Photogrammetry and Remote Sensing 169: 110–124. doi:10.1016/j.isprsjprs.2020.08.025.

Hallaj Z, Bijani M, Karamidehkordi E, Yousefpour R & Yousefzadeh H (2024) Forest land use change effects on biodiversity ecosystem services and human well-being: A systematic analysis. Environmental and Sustainability Indicators 23: 100445. doi:10.1016/j.indic.2024.100445.

Hao Z, Lin L, Post CJ, Mikhailova EA, Li M, Chen Y, Yu K & Liu J (2021) Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN). ISPRS Journal of Pho-

togrammetry and Remote Sensing 178: 112–123. doi:10.1016/j.isprsjprs.2021.06.003.

He L, Zhou Y, Liu L, Zhang Y & Ma J (2025) Research and application of deep learning object detection methods for forest fire smoke recognition. Scientific reports 15: 16328. doi:10.1038/s41598-025-98086-w.

Hua W, Hua C, Zhang S, Qiu T, Jiang X, Li B & Chen B (2025) A Dynamic Model for Estimating Forest Carbon Storage: Application in Wuyishan Forests. Forests 16: 1758. doi:10.3390/f16121758.

Huo Z, Fang L, Chu Y, Dang S, Yang J, Li L, Li X, Ren S, Chen J, Peng Y, Wang G & Wang Q (2026) Precise urban tree species identification and biomass estimation using UAV–Handheld LiDAR Synergy and YOLOv11 deep learning. International Journal of Applied Earth Observation and Geoinformation 146: 105049. doi:10.1016/j.jag.2025.105049.

Hurteau MD (2021) The role of forests in the carbon cycle and in climate change: Climate change. Elsevier, pp. 561–579. doi:10.1016/B978-0-12-821575-3.00027-X.

Jakubowski MK, Li W, Guo Q & Kelly M (2013) Delineating individual trees from lidar data: A comparison of vector- and raster-based segmentation approaches. Remote Sensing 5: 4163–4186. doi:10.3390/rs5094163.

Jegham N, Koh CY, Abdelatti M & Hendawi A (2024) YOLO Evolution: A comprehensive benchmark and architectural review of YOLOv12, YOLO11, and their previous versions. arXiv preprint arXiv:2411.00201. doi:10.13140/RG.2.2.15952.83201.

Jucker T, Fischer FJ, Chave J, Coomes DA, Caspersen J, Ali A, Panzou GJL, Feldpausch TR, Falster D, Usoltsev VA, Adu-Bredu S, Alves LF, Aminpour M, Angoboy IB, Anten NPR, Antin C, Askari Y, Muñoz R, Ayyappan N, Balvanera P, Banin L, Barbier N, Battles JJ, Beeckman H, Bocko YE, Bond-Lamberty B, Bongers F, Bowers S, Brade T, van Breugel M, Chantrain A, Chaudhary R, Dai J, Dalponte M, Dimobe K, Domec JC, Doucet JL, Duursma RA, Enríquez M, van Ewijk KY, Farfán-Rios W, Fayolle A, Forni E, Forrester DI, Gilani H, Godlee JL, Gourlet-Fleury S, Haeni M, Hall JS, He JK, Hemp A, Hernández-Stefanoni JL, Higgins SI, Holdaway RJ, Hussain K, Hutley LB, Ichie T, Iida Y, Jiang HS, Joshi PR, Kaboli H, Kazempour Larsary M, Kenzo T, Kloeppel BD, Kohyama T, Kunwar S, Kuyah S, Kvasnica J, Lin S, Lines ER, Liu H, Lorimer C, Loumeto JJ, Malhi Y, Marshall PL, Mattsson E, Matula R, Meave JA, Mensah S, Mi X, Momo S, Moncrieff GR, Mora F, Nissanka SP, O'Hara KL, Pearce S, Pelissier R, Peri PL, Ploton P, Poorter L, Javanmiri Pour M, Pourbabaei H, Dupuy-Rada JM, Ribeiro SC, Ryan C, Sanaei A, Sanger J, Schlund M, Sellan G, Shenkin A, Sonké B, Sterck FJ, Svátek M, Takagi K, Trugman AT, Ullah F, Vadeboncoeur MA, Valipour A, Vanderwel MC, Vovides AG, Wang W, Wang LQ, Wirth C, Woods M, Xiang W, de Aquino Ximenes F, Xu Y, Yamada T & Zavala MA (2022) Tallo: A global tree allometry and crown architecture database. Global Change Biology 28: 5254–5268. doi:10.1111/gcb.16302.

Jurjević L, Liang X, Gašparović M & Balenović I (2020) Is field-measured tree height as reliable as believed – Part II, A comparison study of tree height estimates from conventional field measurement and low-cost close-range remote sensing in a deciduous forest. ISPRS Journal of Photogrammetry and Remote Sensing 169: 227–241. doi:10.1016/j.isprsjprs.2020.09.014.

Kaartinen H, Hyyppä J, Yu X, Vastaranta M, Hyyppä H, Kukko A, Holopainen M, Heipke C, Hirschmugl M, Morsdorf F, Næsset E, Pitkänen J, Popescu S, Solberg S, Wolf BM & Wu J-C (2012) An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sensing 4: 950–974. doi:10.3390/rs4040950.

Kim J-S, Sung S-M, Back K-S & Lee Y-S (2024) Accuracy assessment of advanced laser scanner technologies for forest survey based on three-dimensional point cloud data. Sustainability 16: 10636. doi:10.3390/su162310636.

Laurin GV, Ding J, Disney M, Bartholomeus H, Herold M, Papale D & Valentini R (2019) Tree height in tropical forest as measured by different ground, proximal, and remote sensing instruments, and impacts on above ground biomass estimates. International Journal of Applied Earth Observation and Geoinformation 82: 101899. doi:10.1016/j.jag.2019.101899.

Liu Y, Zhang A & Gao P (2025) From crown detection to boundary segmentation: Advancing forest analytics with enhanced YOLO model and airborne LiDAR point clouds. Forests 16: 248. doi:10.3390/f16020248.

Lou X, Huang Y, Fang L, Huang S, Gao H, Yang L, Weng Y & Hung I-Ku (2022) Measuring loblolly pine crowns with drone imagery through deep learning. Journal of Forestry Research 33: 227–238. doi:10.1007/s11676-021-01328-6.

Luo T, Rao S, Ma W, Song Q, Cao Z, Zhang H, Xie J, Wen X, Gao W, Chen Q, Yun J & Wu D (2024) YOLOTree-individual tree spatial positioning and crown volume calculation using UAV-RGB imagery and LiDAR data. Forests 15: 1375. doi:-doi:10.3390/f15081375.

Ma J, Ushiku Y & Sagara M (2022) The effect of improving annotation quality on object detection datasets: A Preliminary Study: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 4849–4858.

Marcello J, Spínola M, Albors L, Marqués F, Rodríguez-Esparragón D & Eugenio F (2024) Performance of individual tree segmentation algorithms in forest ecosystems using UAV LiDAR data. Drones 8: 772. doi:10.3390/drones8120772.

Mendes PAS, Coimbra AP & de Almeida AT (2023) Forest vegetation detection using deep learning object detection models. Forests 14: 1787. doi:10.3390/f14091787.

Mo J, Lan Y, Yang D, Wen F, Qiu H, Chen X & Deng X (2021) Deep learning-based instance segmentation method of litchi canopy from UAV-acquired images. Remote Sensing 13: 3919. doi:10.3390/rs13193919.

Nikitina (2024) Automatic segmentation of tree crowns in Pine forests using mask r-cnn on rgb imagery from uavs. Forest Science Issues 7: 1–18. doi:10.31509/2658-607x-202473-150.

Pandey S, Chen KF & Dam EB (2023) Comprehensive multimodal segmentation in medical imaging: Combining YOLOv8 with SAM and HQ-SAM Models: 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), pp. 2584–2590.

Pranata IMD, Darma IWAS, Sandhiyasa IMS & Wiguna IKAG (2023) Strawberry disease detection based on YOLOv8 and K-Fold cross-validation. Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi) 11: 199–210. doi:10.24843/jim.2023.v11.i03.p06.

Prousalidis K, Bourou S, Velivassaki T-H, Voulkidis A, Zachariadi A & Zachariadis V (2024) Olive tree segmentation from UAV imagery. Drones 8: 408. doi:10.3390/drones8080408.

Psistaki K, Tsantopoulos G & Paschalidou AK (2024) An overview of the role of forests in climate change mitigation. Sustainability 16: 6089. doi:10.3390/su16146089.

Que H, Gao H, Shan W, Liu M, An J, Deng F, Feng S, Yang X & Mu L (2026) FM-SAM: individual tree crown delineation and classification based on Segmentation Anything Model (SAM) and YOLOv10 in UAV imagery for forest monitoring. Computers and Electronics in Agriculture 240: 111162. doi:10.1016/j.compag.2025.111162.

Rahman AU, Heinaro E, Ahishali M & Junttila S (2025) Dual-task learning for dead tree detection and segmentation with hybrid self-attention U-Nets in aerial imagery. International Journal of Applied Earth Observation and Geoinformation 144: 104851. doi:10.1016/j.jag.2025.104851.

Ramos LT & Sappa AD (2025) A decade of you only look once (YOLO) for object detection: A review. IEEE Access 13: 192747–192794. doi:10.48550/arXiv.2504.18586.

Rato HMYD, Husni EM & Safitri GR (2025) Tree detection and delineation on aerial imagery of indonesia's forestry using YOLO11 deep learning architecture: 2025 International Seminar on Intelligent Technology and Its Applications (ISITIA), pp. 1–6. doi:10.1109/ISITIA66279.2025.11137498.

Reisi Gahrouei O, Côté J-F, Bournival P, Giguère P & Béland M (2024) Comparison of deep and machine learning approaches for quebec tree species classification using a combination of multispectral and LiDAR data. Canadian Journal of Remote Sensing 50: 2359433. doi:10.1080/07038992.2024.2359433.

Saliu IS, Satyanarayana B, Fisol MAB, Wolswijk G, Decannière C, Lucas R, Otero V & Dahdouh-Guebas F (2021) An accuracy analysis of mangrove tree height mensuration using forestry techniques, hypsometers and UAVs. Estuarine, Coastal and Shelf Science 248: 106971. doi:10.1016/j.ecss.2020.106971.

Satama-Bermeo G, Lopez-Guede JM, Rahebi J, Teso-Fz-Betoño D, Boyano A & Akizu-Gardoki O (2025) PRISMA review: Drones and AI in inventory creation of signage. Drones 9: 221–261. doi:10.3390/drones9030221.

Selvaraj MG, Vergara A, Montenegro F, Ruiz HA, Safari N, Raymaekers D, Ocimati W, Ntamwira J, Tits L, Omondi AB & Blomme Guy (2020) Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin. ISPRS Journal of Photogrammetry and Remote Sensing 169: 110–124. doi:10.1016/j.isprsjprs.2020.08.025.

Sengun E, Aksoy S, Sertel E & Fransson JE (2025) Comparative analysis of YOLOv8 and YOLOv11 on tree detection using UAV RGB and laser scanning data. ISPRS annals of the photogrammetry, remote sensing and spatial information sciences 10: 173–179. doi:10.5194/isprs-annals-X-2-W2-2025-173-2025.

Suarez-Fernandez GE, Martinez-Sanchez J & Arias P (2025) Enhancing carbon stock estimation in forests: Integrating multi-data predictors with random forest method. Ecological Informatics 86: 102997. doi:10.1016/j.ecoinf.2025.102997.

Teng M, Ouaknine A, Laliberté E, Bengio Y, Rolnick D & Larochelle H (2025) Assessing SAM for Tree Crown Instance Segmentation from Drone Imagery. arXiv preprint arXiv:2503.20199. doi:10.48550/arXiv.2503.20199.

Ullah A (2025) Community and institutional drivers of deforestation, environmental impacts, and extension interventions for forest management in the Hindu Kush Himalaya. Land Degradation & Development 36: 3575–3586. doi: 10.1002/ldr.5586.

Wang L, Wei F, Tagesson T, Fang Z & Svenning J-C (2025) Transforming forest management

through rewilding: Enhancing biodiversity, resilience, and biosphere sustainability under global change. One Earth 8: 101195. doi:10.1016/j.oneear.2025.101195.

Wołk K & Tatara MS (2024) A review of semantic segmentation and instance segmentation techniques in forestry using LiDAR and imagery data. Electronics 13: 4139. doi:10.3390/electronics13204139.

Windrim L & Bryson M (2020) Detection, segmentation, and model fitting of individual tree stems from airborne laser scanning of forests using deep learning. Remote Sensing 12: 1469. doi:10.3390/rs12091469.

Xin H, Zhang R, Zhang L, Xu H, Yu X, Gou X & Gao Z (2025) Efficient estimation of plant species diversity in desert regions using UAV-based quadrats and advanced machine learning techniques. Journal of Environmental Management 385: 125614. doi:10.1016/j.jenvman.2025.125614.

Yin D & Wang L (2019) Individual mangrove tree measurement using UAV-based LiDAR data: Possibilities and challenges. Remote Sensing of Environment 223: 34–49. doi:10.1016/j.rse.2018.12.034.

Zhao H, Morgenroth J, Pearse G & Schindler J (2023) A systematic review of individual tree crown detection and delineation with convolutional neural networks (CNN). Current Forestry Reports 9: 149–170. doi:10.1007/s40725-023-00184-3.

Zheng J, Fu H, Li W, Wu W, Yu L, Yuan S, Tao WYW, Pang TK & Kanniah KD (2021) Growing status observation for oil palm trees using Unmanned Aerial Vehicle (UAV) images. ISPRS Journal of Photogrammetry and Remote Sensing 173: 95–121. doi:10.1016/j.isprsjprs.2021.01.008.

Zhong H, Zhang Z, Liu H, Wu J & Lin W (2024) Individual tree species identification for complex coniferous and broad-leaved mixed forests based on deep learning combined with UAV LiDAR data and RGB images. Forests 15: 293. doi:10.3390/f15020293.