

Manual Annotations Meet Fine-Tuned Foundation Models: A Comparison on Tree Crown Segmentation Task

Rewanth Ravindran *, Janik Steier, Samer Karam, Dorota Iwaszczuk

Technical University of Darmstadt, Dept. of Civil and Environmental Engineering Sciences,
Remote Sensing and Image Analysis, Darmstadt, Germany
(rewanth.ravindran, janik.steier, samer.karam, dorota.iwaszczuk)@tu-darmstadt.de

KEYWORDS: low-rank adaptation (LoRA), individual tree crown (ITC) delineation, segment-anything (SAM), foundational model, BAMFORESTS dataset

ABSTRACT:

Accurate segmentation of individual tree crowns (ITCs) from remote-sensing imagery is essential for forest monitoring and ecological analysis, yet remains challenging due to overlapping canopies and structural variability. The Segment Anything Model (SAM) shows strong generalization capabilities but requires effective prompting and domain adaptation for remote sensing applications. In this study, we investigate a lightweight fine-tuning strategy using Low-Rank Adaptation (LoRA) to adapt SAM for ITC segmentation on the BAMFORESTS dataset. The impact of different prompting strategies is evaluated, including manually annotated point and bounding box prompts, as well as automatically generated bounding boxes derived from a pre-trained tree detector. SAM is fine-tuned with instance-level ITC masks, enabling prompt-aware segmentation of multiple tree crowns per image. Performance is assessed before and after fine-tuning using standard instance segmentation metrics, including IoU and F1-score. Results show that LoRA-based adaptation improves mask delineation and robustness to prompt variability, with bounding box prompts consistently outperforming point-based inputs. Automatically generated prompts enable a fully automated workflow, although their effectiveness depends on detection quality. Evaluation on an independent validation site with manually annotated ITC labels shows that the fine-tuned LoRA-SAM model achieves performance comparable to manual annotations, while significantly reducing annotation effort. These findings highlight the importance of prompt design in adapting foundation models for remote sensing tasks and demonstrate that parameter-efficient fine-tuning provides a practical pathway toward scalable ITC segmentation.

1. INTRODUCTION

Individual tree crown (ITC) delineation and mapping constitute a fundamental component of modern forest monitoring, enabling detailed analysis, modelling and management of natural, planted and urban forests. At the level of individual trees, accurate crown delineation supports applications such as carbon stock estimation, biodiversity assessment and forest health monitoring, which are increasingly important in the context of climate change and sustainable forest management. By precisely defining crown boundaries, derived metrics such as crown projection area—defined as the vertical projection of the crown onto a horizontal plane—can be estimated, allowing inference of tree diameter, biomass and growth rates (Dalponte et al., 2018; Zhao et al., 2023).

Research into Individual Tree Crown (ITC) segmentation using remote sensing techniques originated in the 1980s, with earliest methods focusing on the identification of crown centroids and the application of watershed transformation algorithms (Weinstein, 2019). With the increasing availability of very high resolution (VHR) datasets, deep learning methods - particularly convolutional neural networks (CNNs) - have become the dominant approach for ITC detection and delineation. Object detection architectures such as RetinaNet and instance segmentation frameworks such as Mask R-CNN and YOLO have demonstrated substantial improvements over traditional image processing techniques (Weinstein, 2019; Minaee et al., 2022). However, despite these advances, the performance of deep learning models remains strongly dependent on the availability of large, high-quality annotated datasets. Precise annotations for ITCs presents

a unique challenge compared to traditional semantic segmentation tasks, primarily due to the morphological complexity and overlapping structures inherent in forest canopies Zhen et al. (2016). This scarcity of high-fidelity ground truth data continues to impede the training and deployment of robust supervised deep learning models for ITC segmentation. In practice, training data for ITC delineation are typically generated through manual annotation of tree crowns using bounding boxes or polygon masks, a process that is labour-intensive, time-consuming and prone to interpretation errors (Steier et al., 2024).

The reliance on manual annotations introduces two critical challenges. First, the limited availability of annotated data can lead to overfitting, particularly given the high variability in tree crown appearance caused by species diversity, health status and management practices (Weinstein et al., 2020a). Second, the same manually generated annotations are commonly used both for training and for evaluating model performance, typically through Intersection over Union (IoU)-based metrics. Any inaccuracies or inconsistencies in these annotations therefore directly propagate into the evaluation, potentially leading to biased or misleading performance estimates. Recent work has highlighted that the quality of reference data fundamentally constrains the reliability of validation results in large-scale tree crown segmentation tasks (Khan et al., 2025). This underscores the need for alternative validation strategies that incorporate independent reference sources, such as LiDAR-derived crowns or official tree inventories.

One promising direction to alleviate the training-data bottleneck is to exploit large pre-trained foundation computer vision models such as the Segment Anything Model (SAM) (Kirillov et al., 2023). SAM is designed as a general-purpose, promptable

* Corresponding author

segmentation model that can segment objects in images either in a zero-shot setting or guided by prompts such as points, boxes, masks or text (Zhu et al., 2025). SAM has demonstrated strong generalization across diverse domains and has been increasingly explored for Earth observation applications, including land cover classification, crop field delineation and urban analysis (Osco et al., 2023). However, despite its flexibility, SAM is designed as a general-purpose model is limited to 3-channel RGB image segmentation, and integrating it with multi- or hyper-spectral data remains an active research challenge (Ravindran et al., 2025). SAM may also not perform optimally when directly applied to domain-specific tasks such as ITC segmentation, particularly in complex forest environments where object boundaries are ambiguous and highly variable (Zhang and Liu, 2023).

To address these limitations, recent research has focused on adapting SAM to remote sensing applications through prompt engineering and model fine-tuning. Prompt-based approaches using bounding boxes or point annotation have been explored to guide segmentation and improve accuracy in complex scenes. At the same time, fine-tuning strategies aim to adapt the model to domain-specific data distributions. However, fully fine-tuning large foundation models such as SAM is computationally expensive due to their high parameter count. This has led to the adoption of parameter-efficient fine-tuning (PEFT) techniques such as Low-Rank Adaptation (LoRA), which introduce a small number of trainable parameters while keeping the original model weights fixed (Zhang et al., 2023). Extensions such as ConvLoRA further enhance this approach by incorporating convolutional structures into the adaptation layers, improving spatial consistency and boundary delineation (Zhong et al., 2024).

Despite these advances, several open challenges remain. SAM is inherently category-agnostic and relies heavily on the availability and quality of input prompts, which limits its applicability for fully automated large-scale ITC segmentation. Moreover, the interplay between prompt quality, model adaptation and segmentation accuracy is not yet fully understood in the context of forest remote sensing. In particular, it remains unclear to what extent parameter-efficient fine-tuning methods can improve segmentation performance compared to prompt-only approaches, and how these improvements translate to independent validation datasets with heterogeneous forest conditions.

Our contribution: In this study, we investigate the potential of LoRA-based fine-tuning of SAM for ITC instance segmentation from very high-resolution digital orthophotos. Experiments are conducted on the BAMFORESTS dataset (Troles et al., 2024) for training and internal evaluation, followed by independent validation on a separate study area using both manually annotated tree crowns and LiDAR-derived reference data. This dual evaluation framework enables a comprehensive assessment of segmentation performance while reducing reliance on potentially biased manual annotations. Our main contributions are:

1. Integration of parameter-efficient fine-tuning (LoRA) with SAM for ITC segmentation, enabling adaptation to domain-specific forest imagery without full model retraining
2. Systematic comparison of prompt-based and fine-tuned segmentation approaches, analysing the influence of prompt quality and model adaptation
3. Dual Validation Framework: combining manual annotations and independent LiDAR-derived reference data to assess segmentation accuracy

Based on these contributions, the study addresses the following research questions:

1. To what extent can LoRA-based fine-tuning improve SAM performance for ITC segmentation compared to prompt-only approaches?
2. How does the quality and type of prompts (e.g. points and bounding boxes) influence segmentation accuracy?
3. How consistent are segmentation results when evaluated against manual annotations versus independent LiDAR-derived reference data?

2. DATASETS

2.1 Training Dataset

The BAMFORESTS dataset published by Troles et al. (2024) was utilised as the primary ground-truth dataset for fine-tuning our LoRA-SAM model. BAMFORESTS is a benchmark dataset of ITC masks consisting of VHR 3-band RGB aerial images with ground sampling distance of 1.6-1.8 cm. This benchmark dataset has complete manually annotated polygon delineations of 27000 ITCs across four areas of interest near Bamberg, Germany. The ITC polygon labels are provided in COCO format, and along with the 12872 tiled imagery, comes with predefined train, validation and test split. The imagery exhibits high spatial variability, mostly including very dense canopy regions, overlapping crowns, and heterogeneous tree species distributions. These characteristics make the dataset particularly suitable for evaluating instance-level segmentation approaches under challenging real-world conditions.

2.2 Independent Validation Dataset

Manual annotations of individual tree crowns will serve as a baseline against which the SAM segmentation predictions will be evaluated. The manual annotations were created using DOPs with a 20 cm resolution and RGB bands of the main municipal cemetery of Frankfurt am Main, Germany. The study site was specifically chosen because of the presence of a forest-like plantation and a publicly accessible tree register (not used in this study) and airborne laser scanning (ALS) data, which facilitates the validation. In contrast to the training dataset (see Figure 1), this validation dataset offers a variety of tree density and types, offering a mix of geometrically planned tree avenues, dense forest-like mixed-species stands, and average to sparse tree stands with visually separate tree crowns, as seen in Figures 1 and 5.

The freely available official ALS 3D point cloud data, as well as the Digital Surface Model (DSM), of the federal state of Hesse, Germany covering the same area was used to derive 2D tree segments as validation reference labels. Specifically, the ALS data acquired during 2020 and 2023 were selected to closely match the acquisition dates of DoPs and tree registers. The tree crown labels were derived using the MatLab LiDAR toolbox, by generating canopy height-model (CHM) from the 3D point clouds, and extracting tree attributes to delineate individual trees and generating 2D ITC masks.

3. METHODOLOGY

3.1 Dataset Preprocessing

The original UAV images of size 2048×2048 pixels were rescaled to 512×512 pixels. This transformation was performed

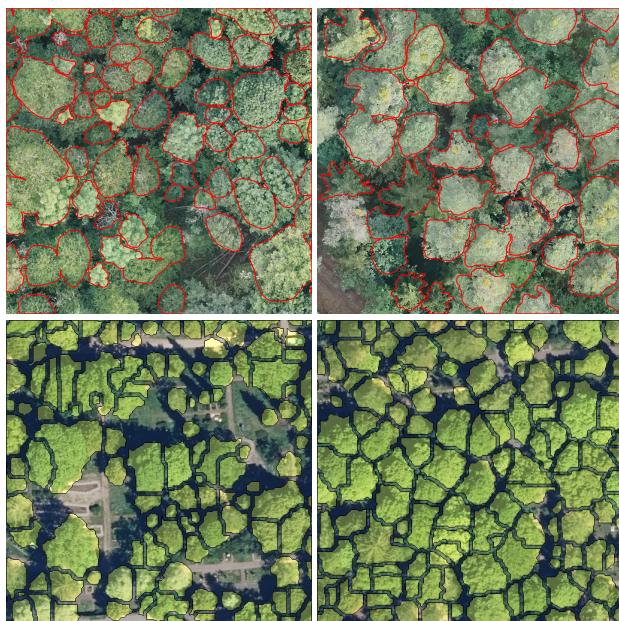


Figure 1. Example images from BAMFORESTS dataset (top) used for fine-tuning SAM and Frankfurt independent validation site (bottom) with ALS data based tree masks

to ensure compatibility with the training pipeline and to reduce computational requirements while preserving sufficient spatial detail for crown delineation. The corresponding COCO annotations were updated accordingly to match the resized images.

Prompt Generation

In addition to the tree labels required for fine-tuning the LoRA layers, SAM requires prompts to guide its mask decoder, which can take the form of point prompts, bounding boxes, or text prompts. Prompts derived from the ground-truth tree-crown annotations of the benchmark dataset were generated to guide the segmentation model during fine-tuning; in this study, we focused on point-based and bounding box prompts, represented in Figure 2.

For point-based prompting, the centroid of each annotated tree crown polygon was computed and used as a foreground point prompt. Each point was assigned a positive label, indicating the presence of a target object. For bounding box prompting, tight axis-aligned bounding boxes were derived from the polygon annotations. These bounding boxes served as spatial priors for the segmentation model. Both types of prompts were used during the LoRA fine-tuning stage. For testing, a subset of ten images was selected, for which manually annotated point and bounding box prompts were generated to evaluate model behaviour under controlled prompting conditions.

Automatic Prompt Generation In addition to ground-truth-based prompts, an automatic bounding box generation approach was employed using the DeepForest Python framework Weinstein et al. (2020b), to ensure the applicability of the method to independent sites. DeepForest is a pre-trained deep learning model designed for tree crown detection in aerial imagery. Bounding boxes were generated using a sliding-window inference strategy, with parameters carefully selected to suit the ITC delineation task. The final configuration was chosen to increase contextual information around tree crowns and to improve detection recall in dense canopy regions. Lower score and IoU

thresholds were applied to retain a larger number of candidate detections, which were subsequently filtered using non-maximum suppression. This approach was applied to both the training dataset and the independent validation dataset. The automatically generated bounding boxes were then used as prompts during inference and evaluation, simulating a fully automated pipeline without reliance on ground-truth annotations.



Figure 2. Types of prompts used - point prompts and bounding box prompts (in cyan) derived from ground truth labels, and automatic bounding box derived using DeepForest method (in red)

3.2 LoRA Fine-Tuning

Low-Rank Adaptation (LoRA) was employed to efficiently adapt the pretrained Segment Anything Model (SAM) to the domain-specific task of individual tree crown (ITC) segmentation. In line with prior work on parameter-efficient fine-tuning of large vision transformers, and subsequent SAM adaptation studies such as SAMed for medical image segmentation (Zhang and Liu, 2023), we adopted a strategy in which the pretrained weights of SAM are kept frozen, while lightweight low-rank updates are introduced into the model. This approach enables efficient adaptation with a significantly reduced number of trainable parameters, which is particularly advantageous when working with limited annotated remote sensing datasets.

The pretrained SAM weights (W_0) were kept frozen, and trainable low-rank matrices were introduced into the self-attention layers of the Vision Transformer (ViT) backbone. Specifically, the Query (Q), Key (K), and Value (V) projection matrices were augmented with trainable low-rank decompositions, such that the effective weight becomes

$$W = W_0 + AB \quad (1)$$

where $A \in R^{d \times r}$ and $B \in R^{r \times k}$ are low-rank matrices, with $r \ll d$. This design follows common practice in recent LoRA-based vision adaptations, where modifying attention projections has been shown to provide a strong balance between expressiveness and parameter efficiency. Prior studies have demonstrated that (a) adapting attention layers is sufficient to capture domain shifts, and (b) restricting updates to low-rank subspaces stabilizes optimisation and reduces overfitting in small datasets—both of which motivated our choice of this implementation over full fine-tuning or heavier adapter-based methods. This formulation enables efficient fine-tuning by significantly reducing the number of trainable parameters compared to full model optimisation.

Implementation: LoRA fine-tuning was conducted for 20 epochs on a reduced dataset consisting of 570 training images, with a batch size of 4 and an initial learning rate of 1×10^{-4} using AdamW optimisation (weight decay 1×10^{-4}). Early

Prompt Method	Before LoRA	After LoRA
Manual Point Prompt	0.23	0.38
Manual Box Prompt	0.37	0.51
Automatic Box Prompt	0.28	0.42

Table 1. IoU metrics for LoRA-SAM ITC segmentation on BAMFORESTS test data.

stopping with a patience of 5 epochs was employed to prevent overfitting. The training loss comprised a combination of binary cross-entropy and Dice loss. The LoRA configuration used a rank of $r = 8$, scaling factor $\alpha = 16$, and dropout of 0.05, providing a compact adaptation of the transformer layers.

3.3 Experiments

To evaluate the influence of prompt design on segmentation performance, our workflow (see Figure 3) consisted of investigating three prompt strategies: (a) ground truth point prompts derived from crown centroids, (b) ground truth bounding box prompts, and (c) automatically generated bounding boxes using a pre-trained tree detection model (DeepForest). For each prompt strategy, two models were evaluated: the baseline SAM without adaptation, using SAM-Geo python package (Wu and Osco, 2023), and the LoRA-adapted SAM. The same prompt inputs were used for both models to ensure a fair comparison of segmentation performance. During training, each image was paired with a corresponding prompt, enabling the model to learn prompt-conditioned instance segmentation. The fine-tuned model was then applied on the independent validation site.

Evaluation Evaluation was conducted separately for each prompt strategy and model variant to analyse the influence of prompt quality and fine-tuning. First, the performance of the DeepForest-based automatic bounding box generation was evaluated by comparing predicted boxes against ground truth annotations. Matching between predicted and ground truth bounding boxes was performed using an Intersection over Union (IoU) threshold of 0.10 to assign supervision targets. Standard detection metrics were computed, including recall, precision and F1-score (Dice-coefficient). Then, segmentation results from each of the three prompt-methods before and after fine-tuning were evaluated using Intersection over Union (mIoU). All of the metrics were computed on a per-image basis to assess the reliability of automatic prompt generation.

4. RESULTS

The quantitative results show that segmentation performance is influenced by both prompting strategy and LoRA-based fine-tuning.

4.1 Training Dataset

Segmentation performance on the BAMFORESTS test subset improved across all prompting strategies after LoRA fine-tuning (Table 1). The largest improvement was observed for manual point prompts (0.23 to 0.38), followed by automatic box prompts (0.28 to 0.42) and manual box prompts (0.37 to 0.51). Manual bounding box prompts achieved the highest IoU both before and after fine-tuning.

Detection performance differed between prompting strategies, as seen in Table 2. Manual prompts achieved higher precision (0.85–0.87) and recall (0.43–0.44) compared to automatic

Metric	Manual Points	Manual BBox	Automatic BBox
Recall	0.44	0.43	0.38
Precision	0.87	0.85	0.67
F1-score	0.58	0.57	0.48

Table 2. Detection performance on BAMFORESTS test subset.

Metric	Manual Annotation	LoRA-SAM
Recall	0.52	0.47
Precision	0.83	0.88
F1-score	0.64	0.61
IoU	0.46	0.48

Table 3. Comparison of detection (precision, recall, F1) and segmentation (IoU) performance between manual annotations and fine-tuned LoRA-SAM on the independent validation site.

bounding boxes (precision 0.67, recall 0.38). Consequently, the F1-score for automatic prompts (0.48) was lower than for manual prompts (0.57–0.58).

4.2 Independent Validation Site

On the independent validation site, LoRA-SAM achieved a recall of 0.47, precision of 0.88, and F1-score of 0.61, compared to manual annotations (recall 0.52, precision 0.83, F1-score 0.64). In terms of segmentation accuracy, LoRA-SAM achieved a higher IoU (0.48) than manual annotations (0.46) (Table 3).

5. DISCUSSION

The results demonstrate that segmentation performance is strongly influenced by both prompting strategy and model adaptation. Addressing the first research question, LoRA fine-tuning consistently improved segmentation accuracy across all prompting strategies, confirming that lightweight adaptation is sufficient to transfer SAM to domain-specific forest imagery. The largest improvements were observed for weaker prompt types, indicating that LoRA enhances the model's ability to capture structural characteristics such as crown geometry and spatial context.

With respect to the second research question, prompt type and quality were found to be critical. Bounding box prompts consistently outperformed point-based prompts, highlighting the importance of spatial constraints for reliable instance delineation. Manual bounding box prompts achieved the highest performance, while automatic prompts enabled a fully automated workflow with moderate loss in accuracy. The reduced performance of automatic prompts reflects their lower detection precision and recall, which propagate into segmentation results.

Qualitative analysis as shown in Figure 4 shows that LoRA-SAM improves boundary delineation and separation of adjacent crowns. In addition, not all automatically generated bounding boxes resulted in segmentation outputs, suggesting that the model is able to reject ambiguous or noisy prompts after fine-tuning.

Regarding the third research question, differences between manual annotations and model outputs were observed on the independent validation site. Manual annotations achieved higher recall, indicating more complete detection of tree instances. However, this did not translate into higher segmentation accuracy, as reflected by lower IoU values.

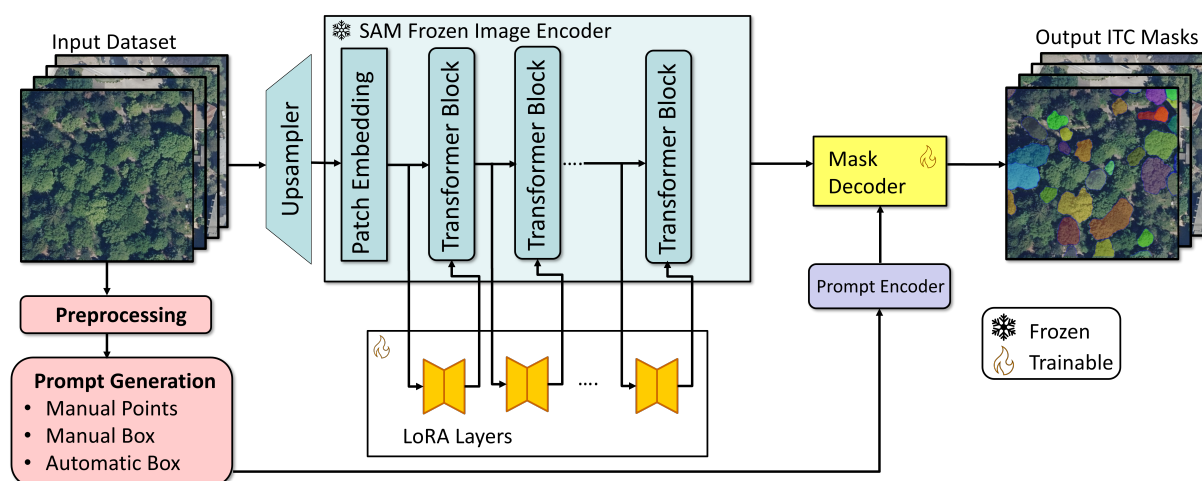


Figure 3. Methodology Workflow depicting the external fine-tuning of Segment Anything (SAM) using Low-Rank Adaptors (LoRA) layers, that modify the transformer blocks in the SAM image encoder. Also depicted are the different prompting techniques used for the SAM mask decoder

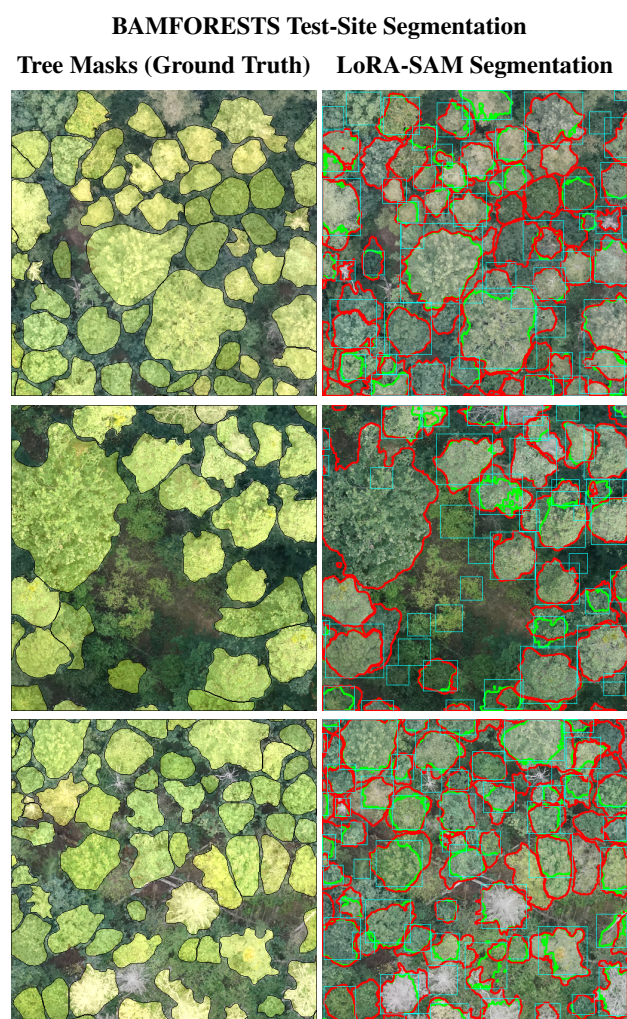


Figure 4. Qualitative comparison of BAMFORESTS ground truth and LoRA-SAM segmentation.

This divergence between F1-score and IoU can be explained by differences in delineation style. Manual annotations tend to produce larger polygons that merge adjacent crowns, improving detection metrics but reducing geometric precision. In contrast, LoRA-SAM produces more spatially constrained segmentations that better follow local canopy structure, leading to improved IoU despite slightly lower recall.

Furthermore, the influence of reference data quality plays a significant role. The LiDAR-derived canopy height model introduces over-segmentation due to multiple local maxima per tree crown, which affects both manual and model-based evaluation. Qualitatively, Figure 5 shows that LoRA-SAM outputs more closely follow the visual structure of the canopy, especially in dense forest patches where crown boundaries are less distinct. Performance varied across tiles, with sparse tree distributions yielding more accurate detection and segmentation, whereas densely packed regions introduced greater ambiguity. Nevertheless, the model demonstrates stable behaviour across varying conditions, indicating good generalisation capability. Additionally, despite a substantial difference in ground sampling distance (2 cm vs. 20 cm), the fine-tuned SAM model produced stable segmentation outputs.

6. CONCLUSION

This study demonstrated that LoRA-based fine-tuning enables efficient adaptation of the Segment Anything Model for individual tree crown segmentation, improving performance across all prompting strategies. Bounding box prompts consistently outperformed point-based prompts, while automatic prompting enabled a scalable workflow with moderate loss in accuracy.

On an independent validation site, LoRA-SAM achieved segmentation performance comparable to manual annotations, while requiring substantially less manual effort. The results further highlight that prompt quality and reference data characteristics remain key factors influencing performance. Overall, the approach provides a practical pathway toward scalable and semi-automated ITC segmentation.

Independent Validation-Site Segmentation

ALS-derived Tree Masks (Ground Truth)

LoRA-SAM Segmentation Results

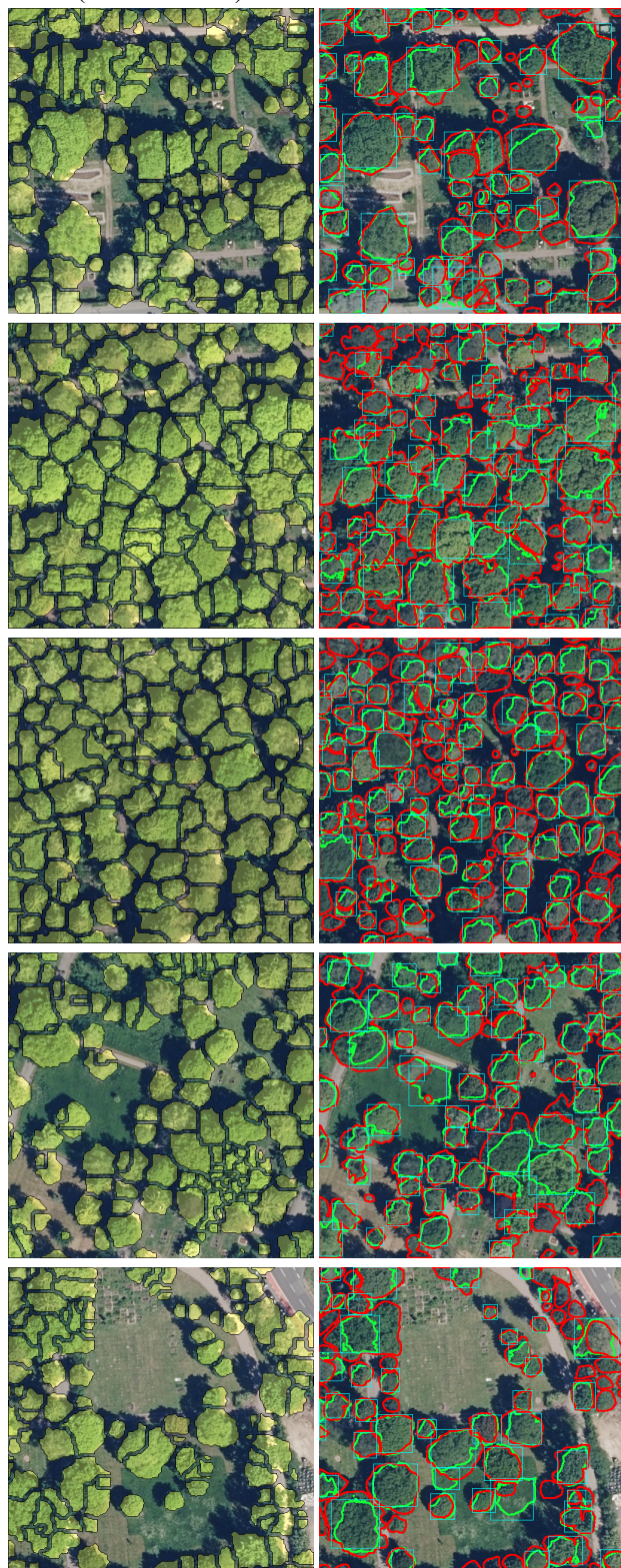


Figure 5. Qualitative comparison of ALS-derived reference tree crowns and LoRA-SAM segmentation outputs (green) in the independent validation site with corresponding bounding box prompts (cyan) and manual annotations (red).

6.1 Outlook

Future work should focus on improving automated prompting strategies and parameter-efficient adaptation methods, such as ConvLoRA Zhong et al. (2024), to enhance segmentation robustness. Training on more diverse datasets and analysing scale effects will further improve generalisation.

Improving reference data quality, for example through refined manual annotations or tree register data, remains essential for reliable evaluation. Finally, extending the framework to newer architectures such as SAM 2 may enable more efficient large-scale forest monitoring.

7. ACKNOWLEDGEMENTS

This study was conducted as part of the cooperative project "WTT GeoKI: For the Collection, Analysis, and Processing of Geodata for Environmental Modeling and Sustainable Solutions," a collaboration between the departments of Remote Sensing Image Analysis and Serious Games at TU Darmstadt. This project is funded by the State of Hesse's program for the European Regional Development Fund (EFRE).

The authors gratefully acknowledge the Hessian State Bureau for Land Management and Geoinformation and the Frankfurt Parks and Green Spaces Department for providing the official geospatial data used in this study.

REFERENCES

- Dalponte, M., Ørka, H., Ene, L., Gobakken, T., Næsset, E., 2018. Tree crown delineation and tree species classification in forest stands using hyperspectral and LiDAR data. *Remote Sensing of Environment*, 213, 292–302.
- Khan, T., Krebs, J., Gupta, S., Renkel, J., Arnold, C., Nölke, N., 2025. Validation challenges in large-scale tree crown segmentations from remote sensing imagery using deep learning: A case study in germany. *EarthArXiv*.
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. *arXiv:2304.02643 [cs]*.
- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., Terzopoulos, D., 2022. Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3523–3542.
- Oscó, L. P., Wu, Q., De Lemos, E. L., Gonçalves, W. N., Ramos, A. P. M., Li, J., Marcato, J., 2023. The Segment Anything Model (SAM) for remote sensing applications: From zero to one shot. *International Journal of Applied Earth Observation and Geoinformation*, 124, 103540.
- Ravindran, R., Treitz, P., Iwaszczuk, D., 2025. Investigation on Dimensionality Reduction Methods for Tree-Crown Segmentation in Hyperspectral Imagery Using Segment Anything Model. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-G-2025, 705–712.
- Steier, J., Goebel, M., Iwaszczuk, D., 2024. Is Your Training Data Really Ground Truth? A Quality Assessment of Manual Annotation for Individual Tree Crown Delineation. *Remote Sensing*, 16(15), 2786.

Troles, J., Schmid, U., Fan, W., Tian, J., 2024. BAMFORESTS: Bamberg Benchmark Forest Dataset of Individual Tree Crowns in Very-High-Resolution UAV Images. *Remote Sensing*, 16(11), 1935.

Weinstein, B. e. a., 2019. Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks. *Remote Sensing*, 11(11), 1309.

Weinstein, B. G., Graves, S. J., Marconi, S., Singh, A., Zare, A., Stewart, D., Bohlman, S. A., White, E. P., 2020a. A benchmark dataset for individual tree crown delineation in co-registered airborne RGB, LiDAR and hyperspectral imagery from the National Ecological Observation Network. Pages: 2020.11.16.385088 Section: New Results.

Weinstein, B. G., Marconi, S., Bohlman, S. A., Zare, A., White, E. P., 2020b. DeepForest: A Python package for RGB deep learning tree crown delineation. *Methods in Ecology and Evolution*, 11(12), 1743–1747.

Wu, Q., Osco, L. P., 2023. samgeo: A Python package for segmenting geospatial data with the Segment Anything Model (SAM). *Journal of Open Source Software*, 8(89), 5663.

Zhang, C., Liu, L., Cui, Y., Huang, G., Lin, W., Yang, Y., Hu, Y., 2023. A Comprehensive Survey on Segment Anything Model for Vision and Beyond. arXiv:2305.08196 [cs].

Zhang, K., Liu, D., 2023. Customized segment anything model for medical image segmentation.

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, 1–22.

Zhen, Z., Quackenbush, L. J., Zhang, L., 2016. Trends in Automatic Individual Tree Crown Detection and Delineation—Evolution of LiDAR Data. *Remote Sensing*, 8(4), 333. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

Zhong, Z., Tang, Z., He, T., Fang, H., Yuan, C., 2024. Convolution meets lora: Parameter efficient finetuning for segment anything model.

Zhu, Y., Locke, W., Yuan, J., Zhang, Y., Ma, Q., Liang, L., 2025. Leveraging SAM 2 and LiDAR for automated individual tree crown delineation: A comparative evaluation of prompting methods. *Information Geography*, 1(2), 100025.