Enhancing Orchard Management with Deep Learning: Tree Segmentation Using Geospatial SAM2 Model and Aerial Imagery

Sarbani Kumar[1], Mohammadreza Narimani[2], Ali Moghimi[2], Alireza Pourreza[2]

1. Department of Electrical and Computer Engineering, University of California, Davis, CA, USA
2. Department of Biological and Agricultural Engineering, University of California, Davis, CA, USA

UC DAVIS COLLEGE OF ENGINEERING

UC DAVIS UNIVERSITY OF CALIFORNIA

Abstract

Agricultural management, particularly in orchard landscapes, is facing increasing challenges due to changing climatic conditions and the need for sustainable practices. In response, this project introduces a cutting-edge approach for orchard management using deep learning to segment individual trees and estimate their areas employing NAIP and Google aerial imagery. Central to this initiative is the implementation of the Geospatial Segment Anything Model 2 (SAM2), an advanced deep learning model specifically adapted for precise tree segmentation. Our methodology builds upon traditional methods through hyperparameter tuning for better tree detection, and next by employing rectangular bounding boxes to delineate each tree, significantly refining the accuracy of the segmentation process. By deploying the SAM2 model, we achieve detailed delineation of tree boundaries and precise area calculations, providing essential data for various agricultural applications. This information is critical for correlating tree area with factors such as yield, fruit and nut production, and biomass estimation. The ability to accurately segment and analyze data at the individual tree level revolutionizes yield predictions, optimizes resource allocation, and fosters sustainable orchard management.

Objectives

1. Develop an Integrated Data Acquisition System: Create a robust codebase that allows for the seamless downloading of region-specific NAIP aerial imagery and Google Maps data for precise orchard analysis.

2. Implement and Optimize the SAM2 Model: Utilize the Geospatial Segment Anything Model 2, enhanced with hyperparameter tuning and rectangular bounding boxes, to accurately segment individual trees from other objects within the imagery.

3. Enhance Analytical Accuracy and Accessibility: Calculate and validate the area of segmented trees against ground truth data, and provide a user-friendly interface on Google Earth Engine to enable real-time performance evaluation and usability by end-users.

Methodology

1. Selection of High-Resolution Imagery for Tree Segmentation

The success of our project hinges on the ability to acquire high-resolution aerial imagery, which is crucial for effectively segmenting and analyzing individual trees in orchard environments. To meet this requirement, we selected a combination of Google Map aerial images and the National Agriculture Imagery Program (NAIP), renowned for their detailed spatial resolution of less than one meter. Additionally, we secured limited access to WorldView-3 satellite imagery from Maxar Technologies, offering an unprecedented resolution of 31 cm. This superior resolution significantly enhances our ability to distinguish individual trees from surrounding objects, providing a critical advantage in our segmentation efforts. Figure 1 displays the WorldView-3 Satellite, underscoring its pivotal role in our data acquisition strategy.

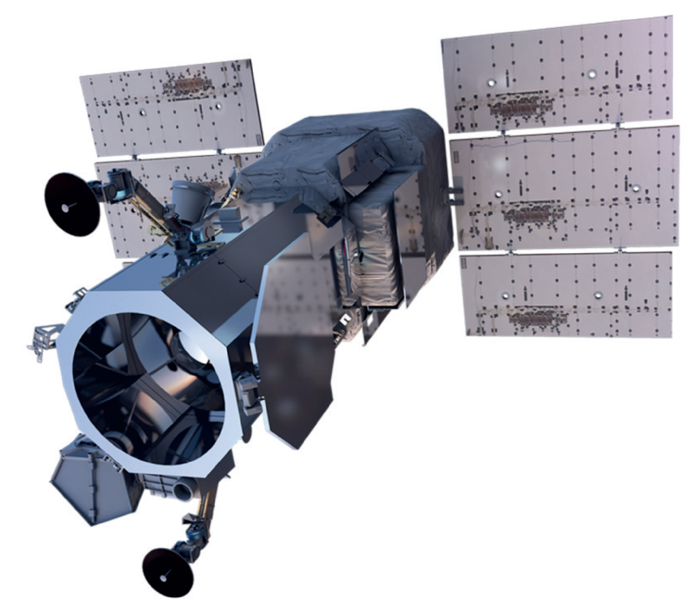

Figure 1. Satellite Image ©2019Maxar Technologies

2. Application of SAM2 for Enhanced Tree Segmentation

The Geospatial Segment Anything Model 2 (SAM2) is essential for accurately segmenting trees using high-resolution aerial imagery. This advanced model not only segments trees but also georeferences the imagery, linking each segment to geographic coordinates for precise applications. SAM2's design includes an Image Encoder, Memory Attention, and Mask Decoder, all crucial for detailed and accurate tree analysis. Figure 2 illustrates the SAM2 architecture.

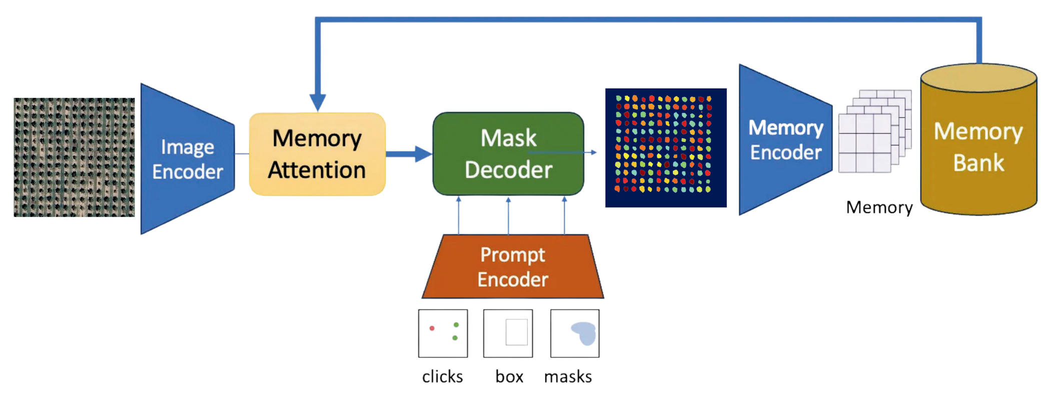

Figure 2. SAM2 Architecture: From Image encoding to mask decoding (Avishek Biswas, 2024).

3. Initial Application and Performance of SAM2

For our project, we deployed the Geospatial Segment Anything Model 2 (SAM2) configured as a high-capability automatic model. This initial setup was specifically tested to assess the model's baseline capability in tree segmentation within orchard environments. Unfortunately, the performance was suboptimal, as SAM2 was only able to detect a single tree, demonstrating significant limitations in handling complex agricultural imagery without tailored adjustments. Figure 3 displays this initial trial, where the left image presents the RGB view of the orchard, and the right image shows the unsatisfactory segmentation result, highlighting the model's initial challenges.

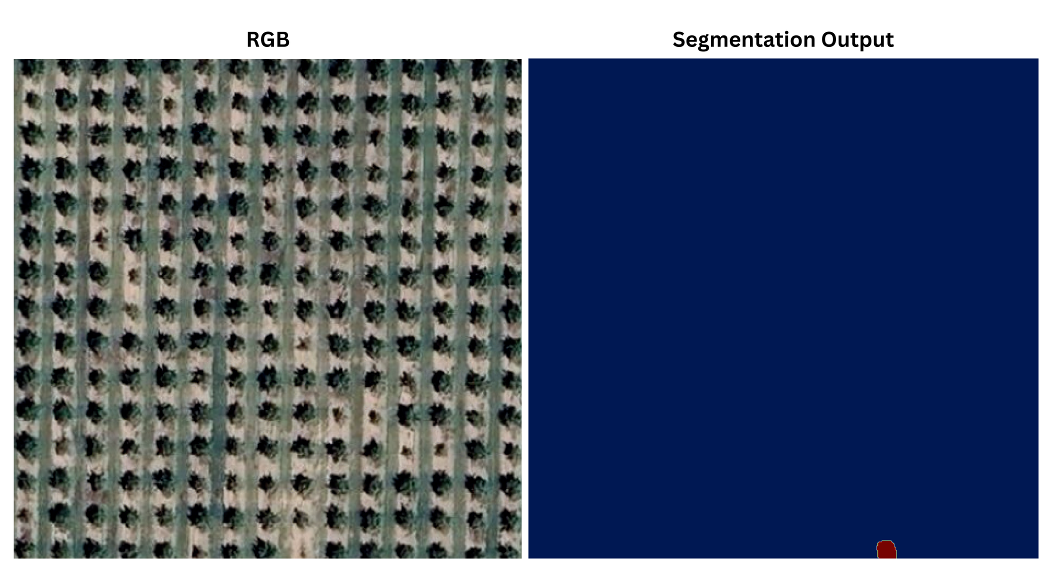

Figure 3. Initial SAM2 Results: Left - RGB image of the orchard; Right - Poor segmentation output.

4. Enhanced Tree Segmentation with Hyperparamter Tuned SAM2

To improve the performance of the Geospatial Segment Anything Model 2 (SAM2), we implemented specific hyperparameter adjustments focused on optimizing segmentation accuracy. These changes significantly enhanced the model's ability to distinguish trees from soil, increasing the reliability of data for agricultural management. Despite notable improvements, some challenges such as misidentifying soil as trees and splitting single trees into multiple segments persisted. Figure 4 shows an RGB image of the orchard on the left and the improved segmentation results on the right after tuning.

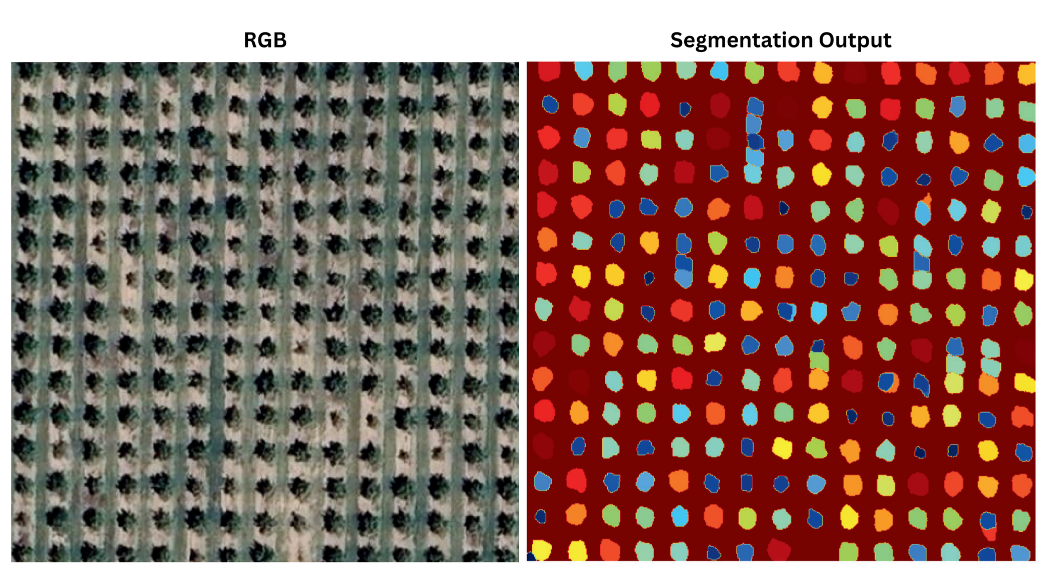

Figure 4. Enhanced SAM2 Segmentation: Left - RGB orchard; Right - Improved results.

5. Improved Tree Segmentation with Defined Bounding Boxes

To significantly improve the precision of tree detection, we utilized the Geospatial Segment Anything Model 2 (SAM2) within a structured 12 by 12 grid of bounding boxes across the orchard. Each bounding box precisely contains one tree, guiding the SAM2 to focus and segment only one object within each box. This setup dramatically enhanced the model's performance by providing clear segmentation boundaries. Figure 5 shows the RGB image of the orchard with the bounding boxes on the left, and the right image illustrates the effective segmentation by SAM2 within these boxes, demonstrating a marked improvement in accuracy.

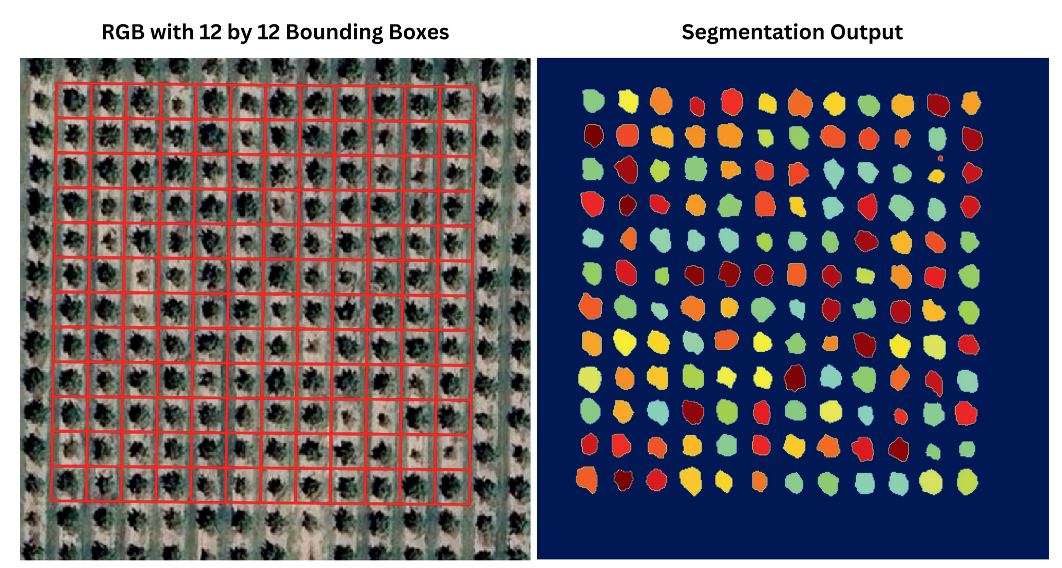

Figure 5. SAM2 Segmentation: Left - Orchard RGB with boxes; Right - Enhanced segmentation.

Results

1. Ground Truth Comparison with Enhanced SAM2 Segmentation

Utilizing the Label Studio package in Python, we efficiently managed the main RGB aerial imagery of the orchard, meticulously brushing pixels identified as trees to establish a ground truth. This manual labeling of canopy pixels is crucial for validating the segmentation accuracy of the Geospatial Segment Anything Model 2 (SAM2). Figure 6 presents a three-part visual representation: the main RGB image on the left, the ground truth with brushed canopy pixels at the center, and the final enhanced SAM2 segmentation result on the right. This comparative display highlights the effectiveness of the SAM2 model in accurately identifying tree coverage.

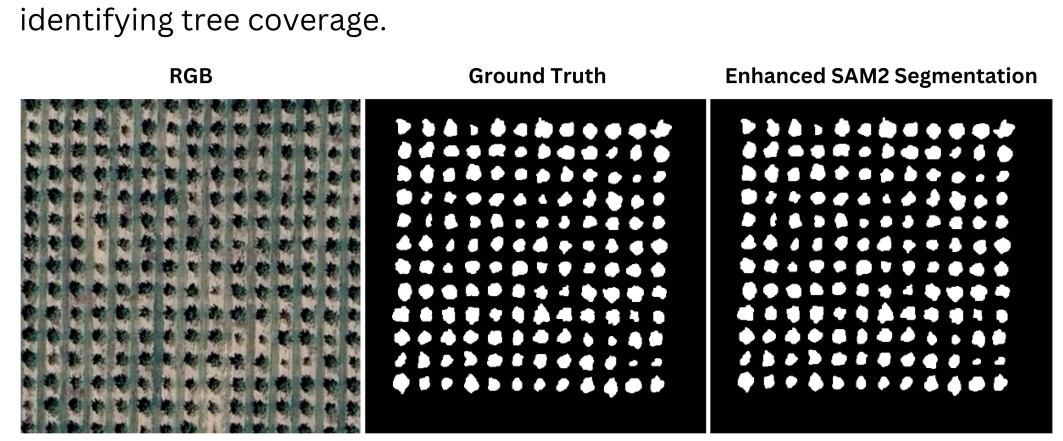

Figure 6. SAM2 Segmentation: Left - RGB image; Center - Ground truth; Right - Enhanced result.

2. Area Estimation Accuracy Demonstrated Through KDE Plots

Utilizing the resolution of aerial imagery and counting pixels for each tree in both the ground truth and SAM2-enhanced segmentation, we calculated the area of each tree. This approach effectively quantifies potential yield and biomass based on area estimations. Figure 7 showcases a Kernel Density Estimate (KDE) plot that compares the tree area distributions as derived from the ground truth and SAM2's estimated segments. This visual representation highlights the alignment and accuracy of SAM2 in replicating the actual tree areas, underscoring its utility in precision agriculture.

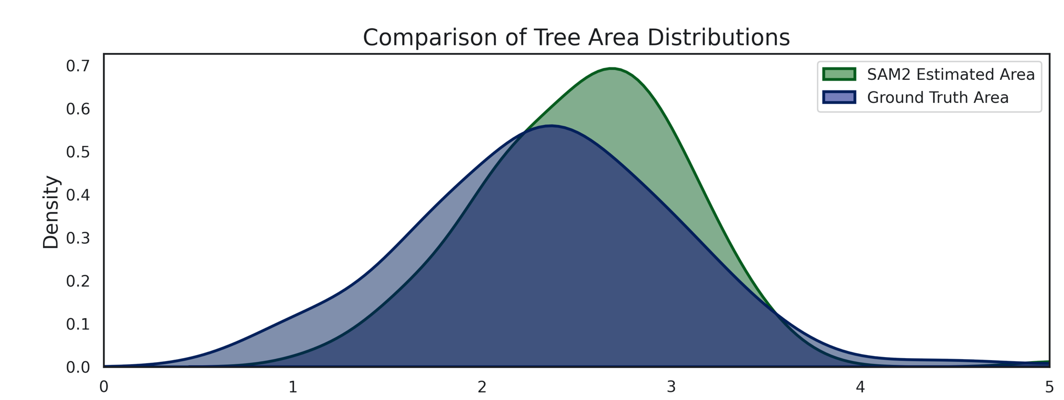

Figure 7. Area Distribution: KDE plot comparing ground truth (blue) to SAM2 estimates (green).

3. Performance Metrics and Segmentation Accuracy of GeoSAM2

We quantitatively evaluated the segmentation performance of the GeoSAM2 model by calculating key metrics, including accuracy, precision, recall, F1 score, and the Intersection over Union (IoU). The results were impressive, with an overall model accuracy of 91.17%. However, the IoU at 61.99% indicates some limitations, such as the model occasionally misclassifying tree shadows as part of the canopy. This is also reflected in the normalized confusion matrix, where a substantial majority of tree pixels were correctly classified, yet some non-tree elements were mistakenly identified as trees. Figure 8 displays these findings, with the confusion matrix on the left showing detailed classification percentages and the metric values table on the right.

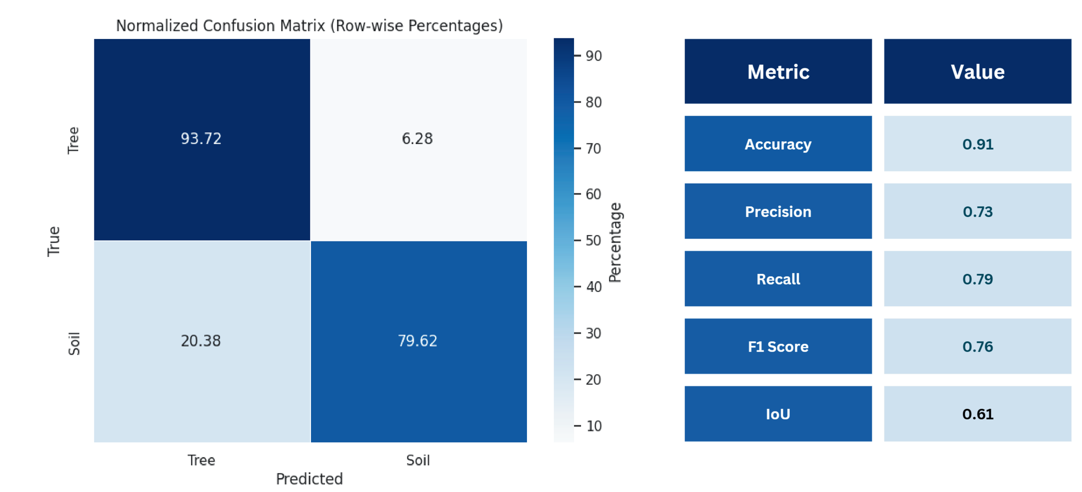

Figure 8. GeoSAM2 Evaluation: Left - Normalized confusion matrix; Right - Performance metrics table.

Interactive Segmentation Tool

We developed an interactive tool on Google Earth Engine that integrates GeoSAM2 model results with Google Maps. This tool features a dynamic interface with a split panel, enabling users to slide between the original aerial imagery and the segmented tree areas. This functionality enhances user engagement and improves understanding of the model's accuracy and effectiveness in real-world applications. It's particularly useful for orchard managers and researchers, providing an intuitive way to assess tree distribution and area in any selected orchard. Figure 9 showcases the tool's interface, illustrating the side-by-side comparison capability.

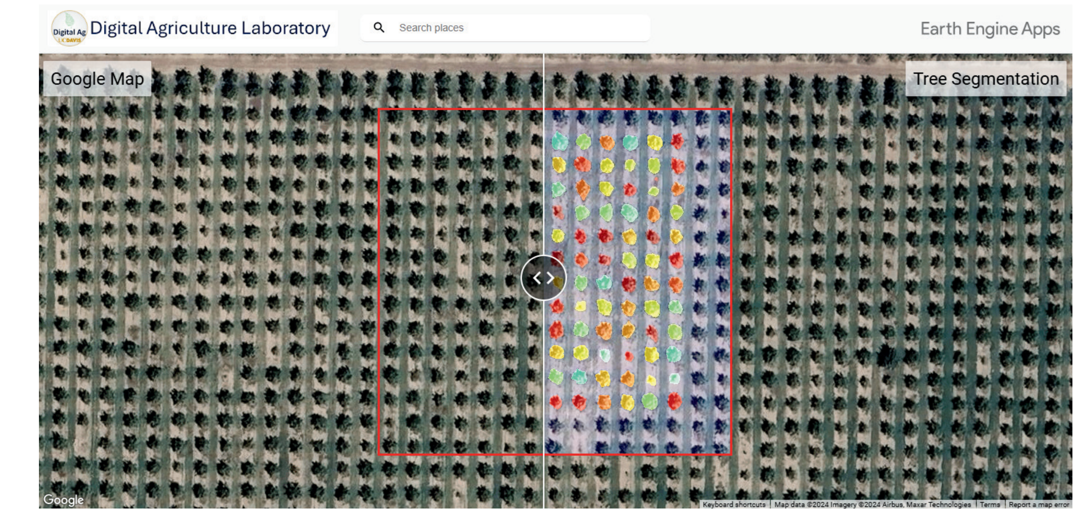

Figure 9. Interactive Tool: GEE dashboard shows original imagery and GeoSAM2 results in an orchard.

Explore Our Tool

bit.ly/GeoSAM2TreeSegmentation

Digital Ag UC DAVIS