Strawberry Canopy Geometric Parameters Estimation and Dry Biomass Prediction from UAV Multispectral Imagery Using Machine Learning Methods

By Caiwang ZHENG,

04/21/2023

UF | IFAS
UNIVERSITY of FLORIDA
Gulf Coast Research and Education Center

@Plant City

SCHOOL OF FOREST, FISHERIES, AND GEOMATICS SCIENCES

UF GEOMATICS

Table of Contents

1. Introduction

In the past decade, **image-based high-throughput phenotyping** technologies have rapidly developed in the field of precision agriculture. Remotely sensed image acquisition and analysis enable the accurate and rapid measurement of plant traits for plant breeding and crop management. However, **modeling strawberry plant canopy geometric and biophysical parameters (e.g., dry biomass) at the individual plant level remains a major challenge.**

Physiological Traits (e.g., chlorophyll, biomass, pigment content, photosynthesis)

Geometric Parameters (e.g., plant height, area, volume)

Yield (fruit/flower number)

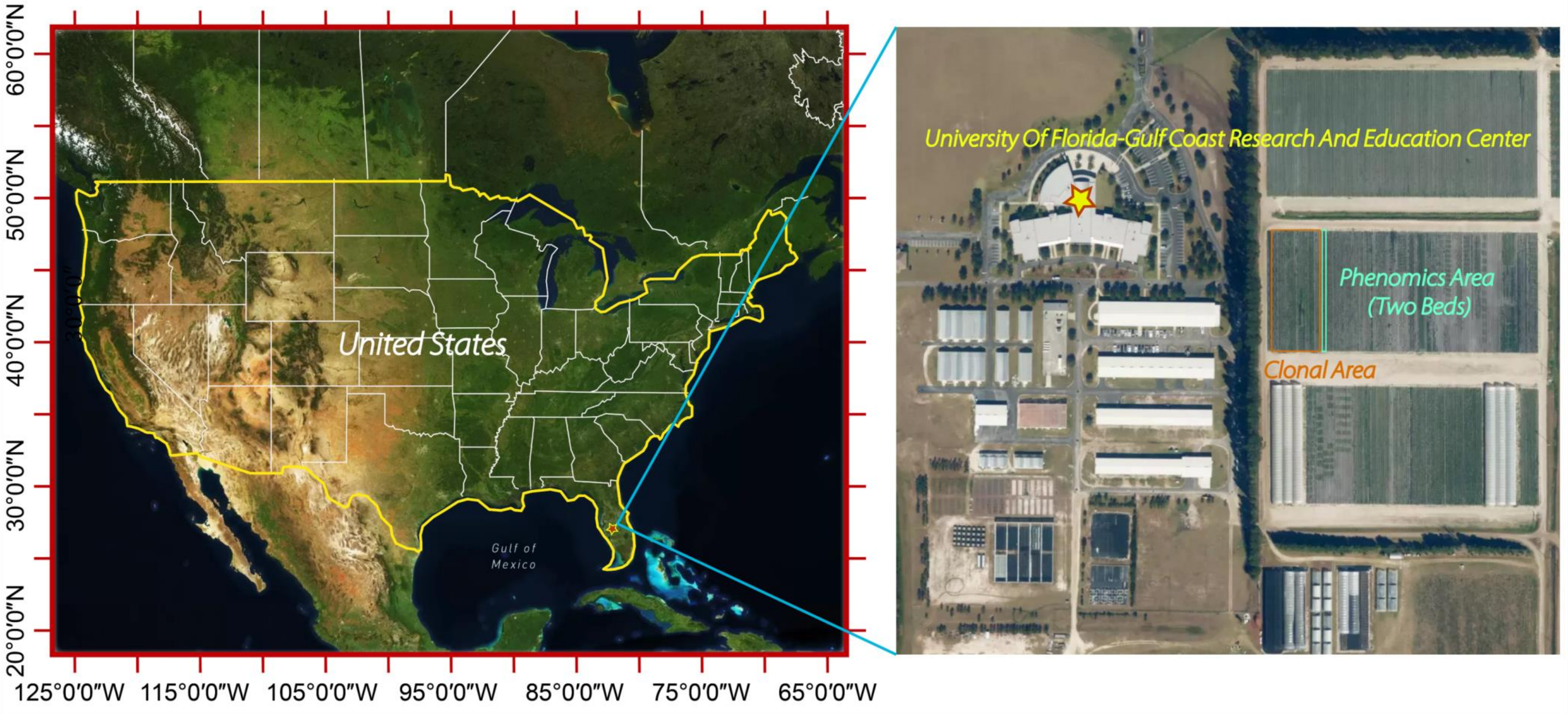
Abiotic/biotic stress indicators (e.g., disease, leaf water)

More traits?

1. Introduction

Phenomics Area: 17 genotypes, ~500 plant samples with in-situ biomass and height measurements.

Biomass Modeling



Biomass Prediction

Clonal Area: more than 2000 strawberry plants with about 400 genotypes/varieties for plant breeding and genetic analysis.

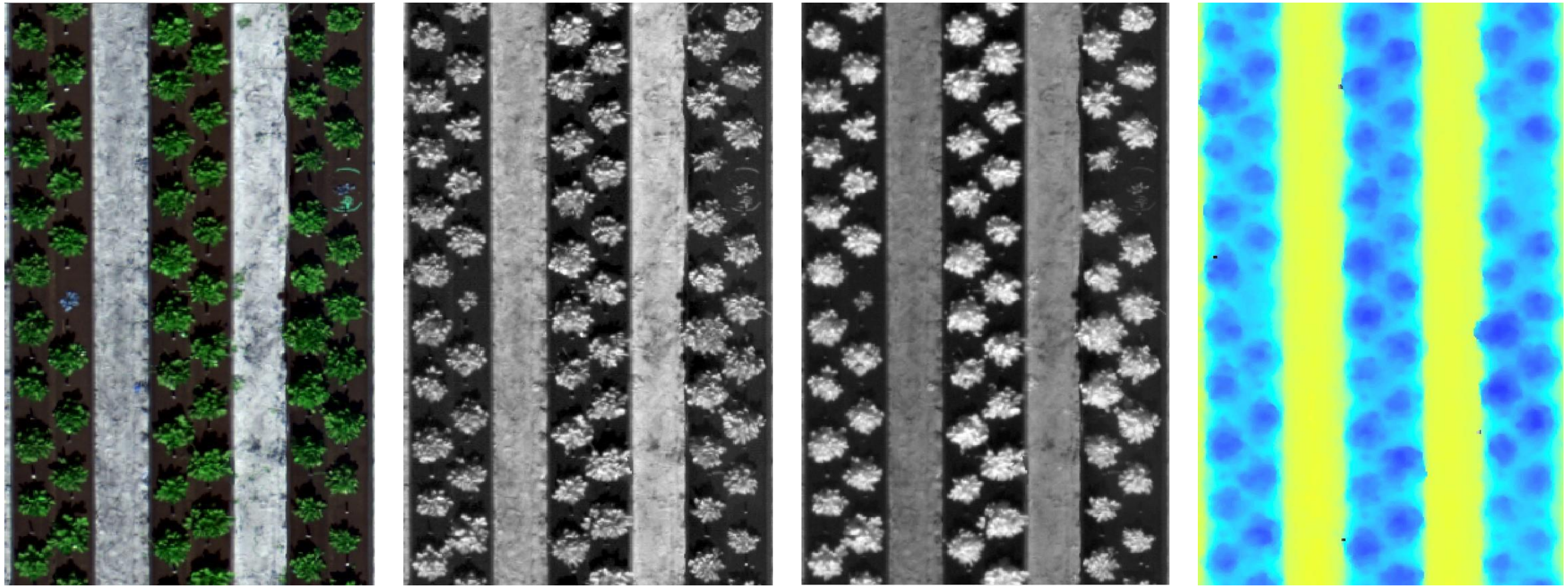
UF IFAS UNIVERSITY of FLORIDA

2. Field Work – Data Collection in 2020-2022 season

Inspire 2 drone mounted with MicaSense Red-Edge M multispectral camera (5 bands)

UAV image data collection and in-situ biomass measurements were performed weekly from November 2020 to March 2022 (about 16 weeks).

3. Orthomosaic Image and DSM generation



RGB NIR RedEdge DSM

Using the **Structure from Motion (SfM)** algorithm, the **Agisoft Metashape** combined the images taken from multiple angles and generated dense 3D point clouds of scene objects (strawberry canopies, soils, and beds). A **digital surface model (DSM)** was created from the SfM-based point cloud, and a **georeferenced orthomosaic image** was generated for each band. The spatial resolution of both DSM and orthomosaic images had **1-cm pixel size**.

UF IFAS UNIVERSITY of FLORIDA

Objective 1: Strawberry Canopy Delineation/Instance Segmentation from the UAV Multispectral Imagery

1). Strawberry Canopy Delineation Methods

Table 1. Strawberry canopy delineation methods summary

Type	Method	Citation	Usage
Computer vision method	Marker Controlled Watershed Algorithm (MCWA)	Abd-Elrahman et al., 2020	Training samples preparing for deep learning models
Deep learning models	Mask R-CNN	He et al., 2017	Strawberry plant instance segmentation performance evaluation and comparison
	YOLACT	Bolya et al., 2019	
	SOLOv2	Wang et al., 2020	
	Mask2Former	Cheng et al., 2022	

1). Marker Controlled Watershed Algorithm (MCWA)

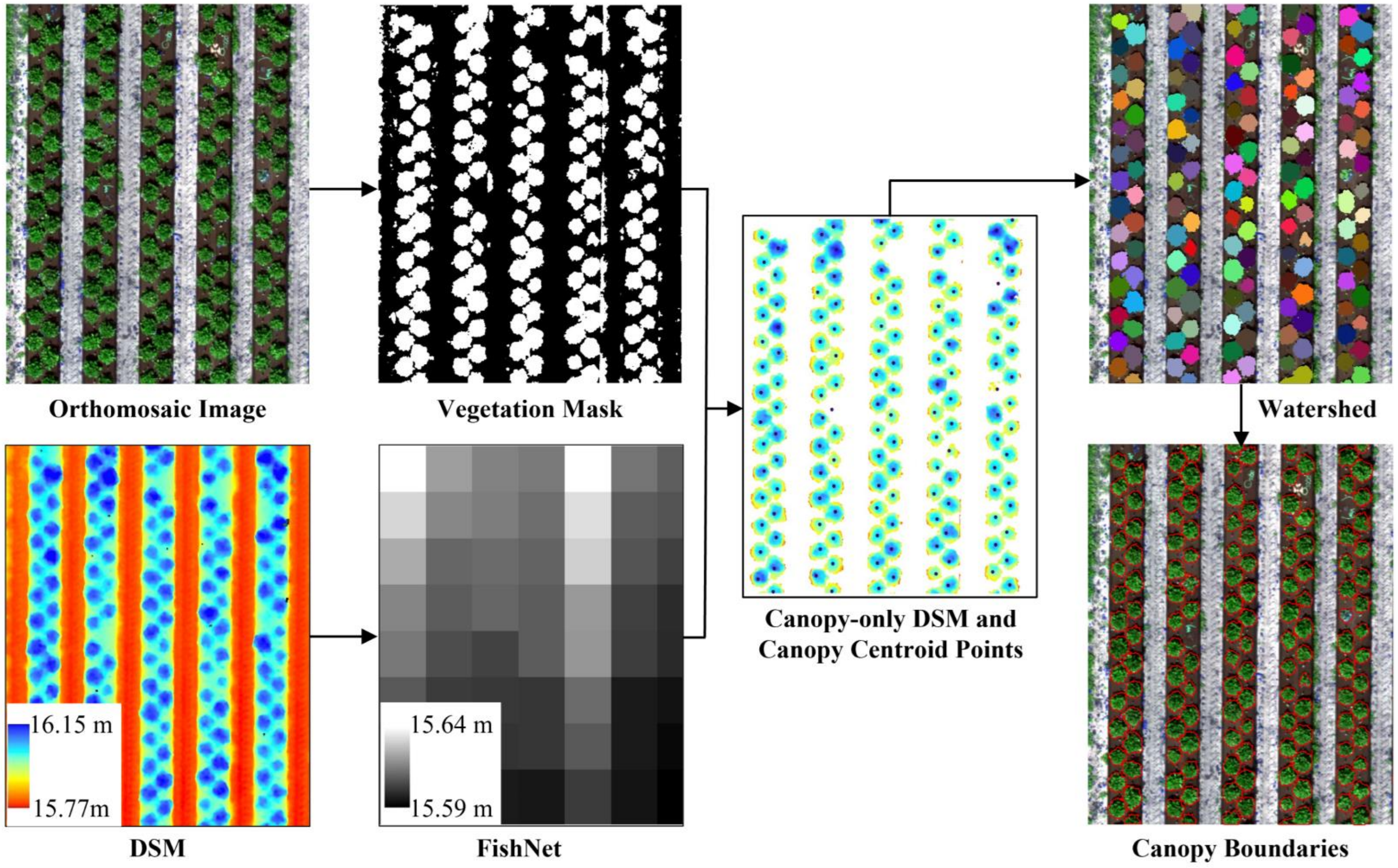


Figure 1. Illustration of the automatic canopy delineation process

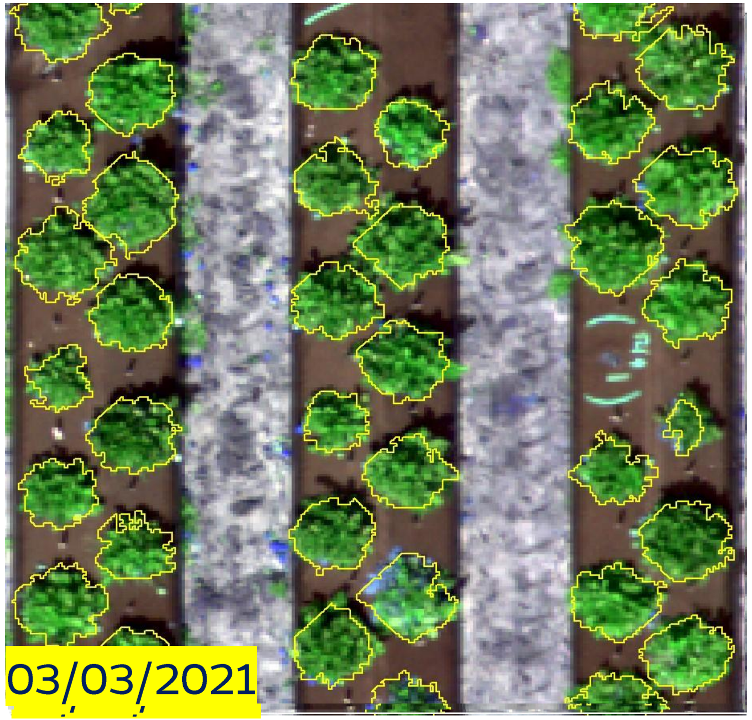
The R script was used to perform the Marker Controlled Watershed Algorithm (MCWA) was adopted for canopy delineation.

1) **Vegetation mask** generation using the NDVI through the **OSTU threshold method**;
2) The 1 m square cells (**Fishnets**) were generated to set up the local elevation threshold to filter out the soil and grass objects;
3) Based on the **canopy centroid points and canopy-only DSM**, the **MCWA** was used to generate canopy boundaries.

2).Training Samples Preparation for Deep Learning Models

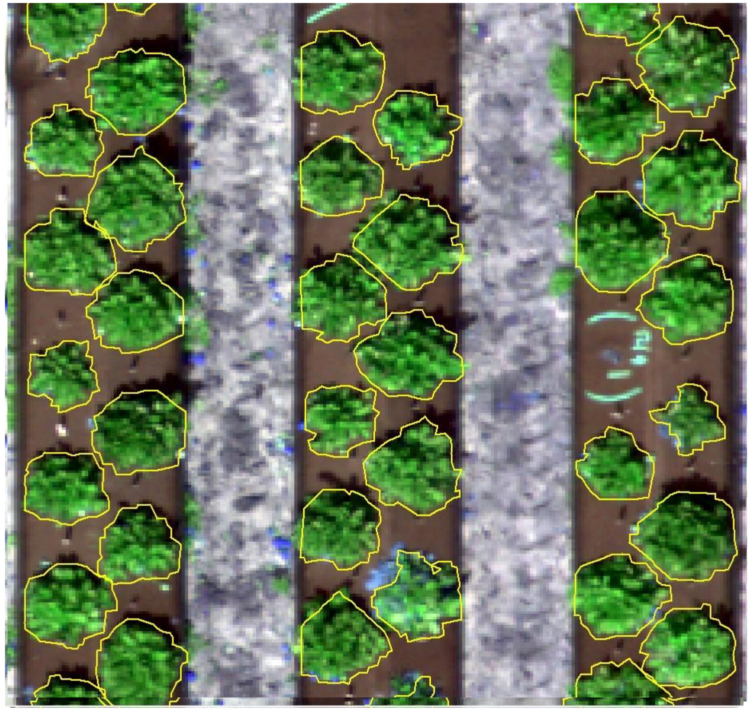
Table 2. UAV image dataset for deep learning model training and test

Season	Number of data collections	Usage
2020-2021	13	Model training and validation
2021-2022	28	Only for model test



MCWA-generated training samples

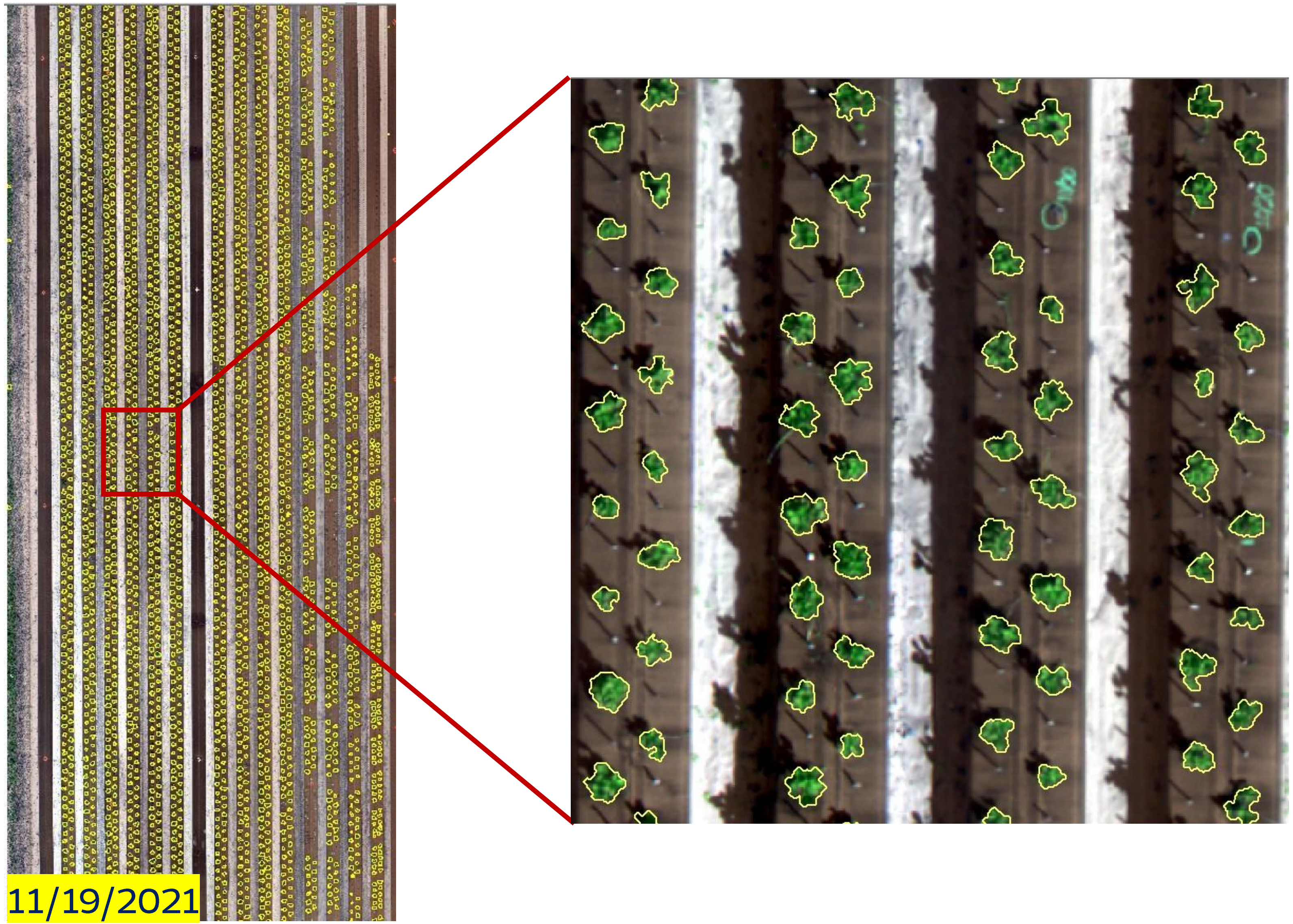
1 cm buffer analysis



Smoothed training samples

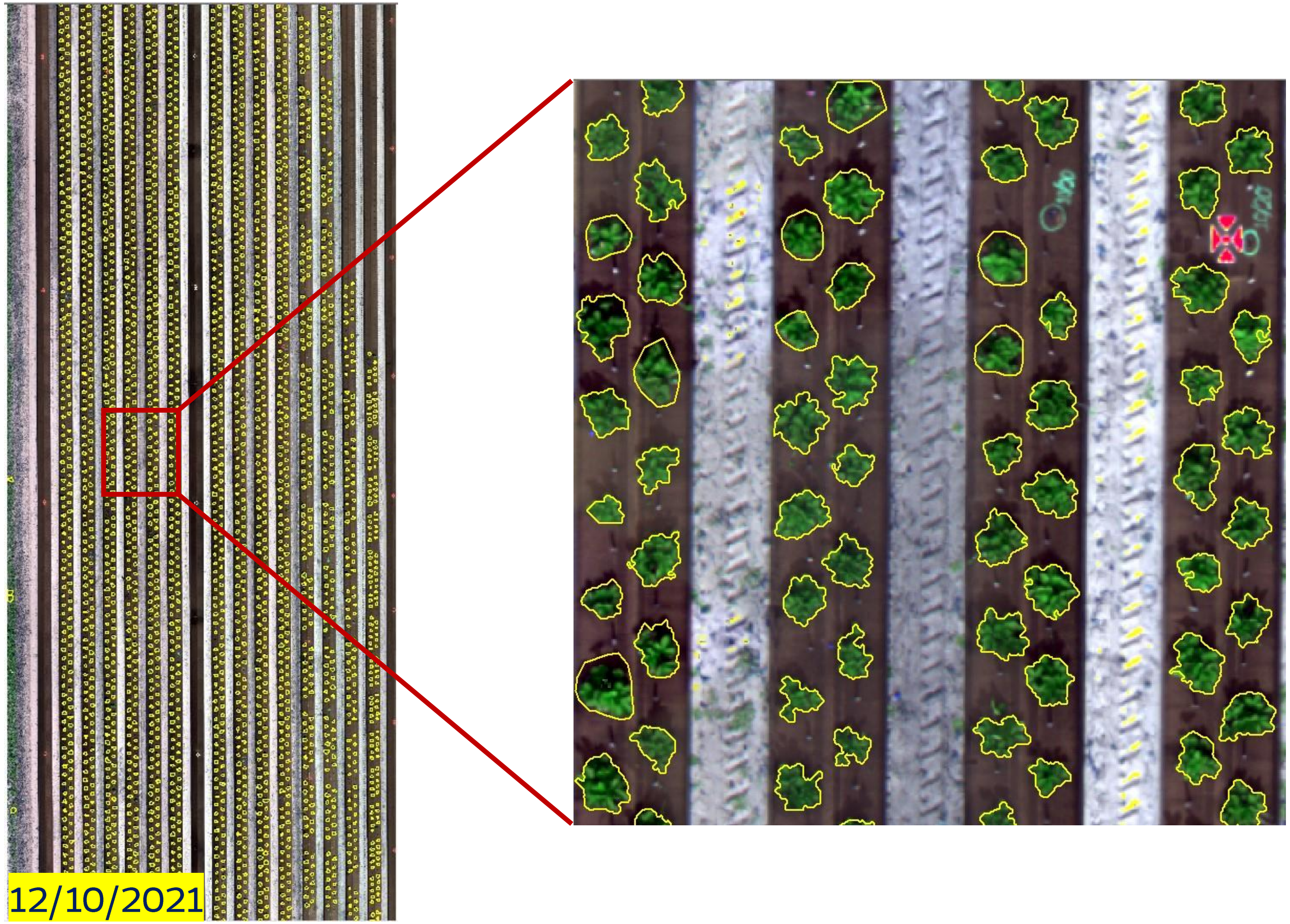
A total of 354 images with 512×512-pixel tiles were used for deep learning model training.

3). Strawberry Canopy Delineation using Deep Learning



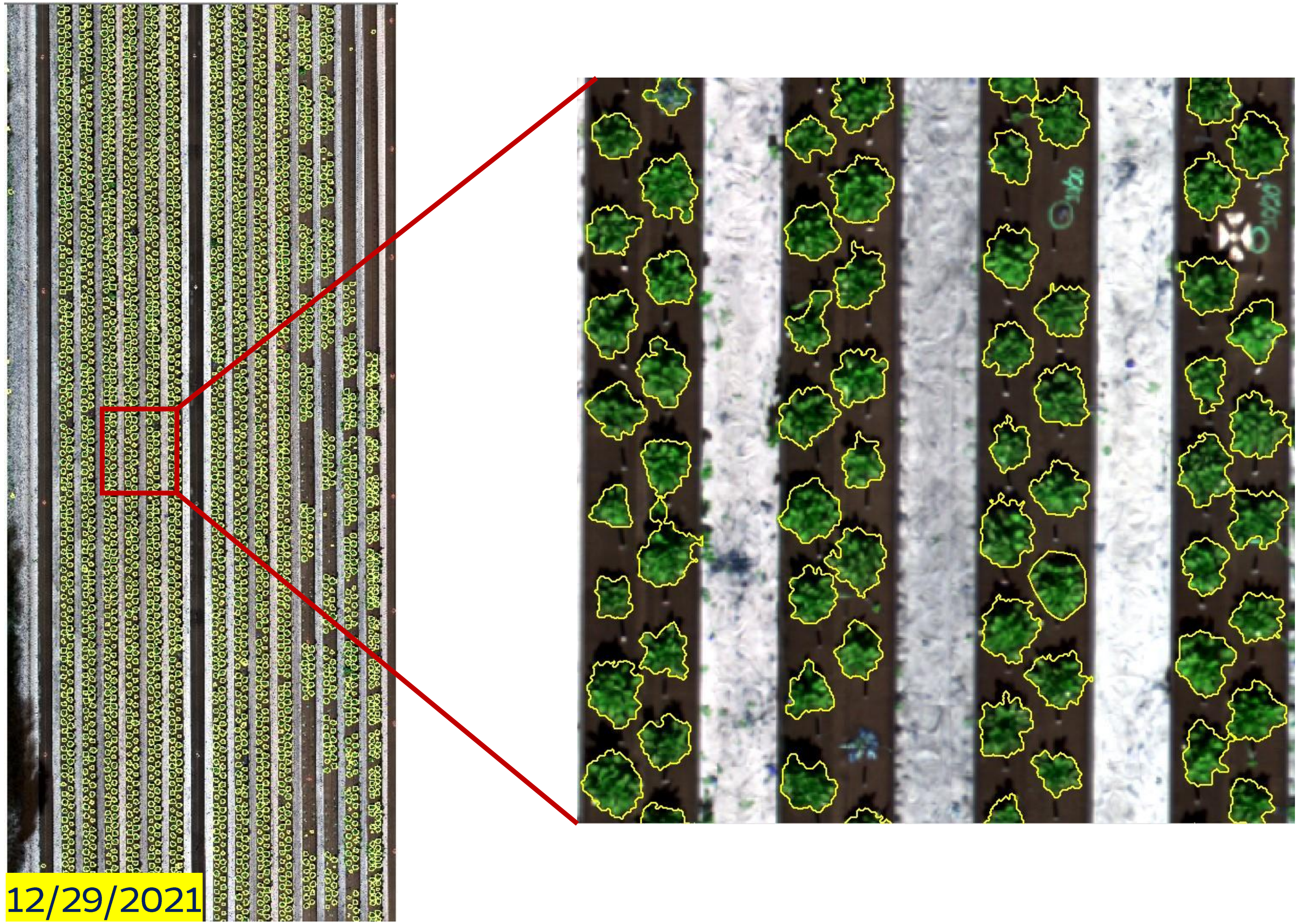
UF | IFAS
UNIVERSITY of FLORIDA

3). Strawberry Canopy Delineation using Deep Learning

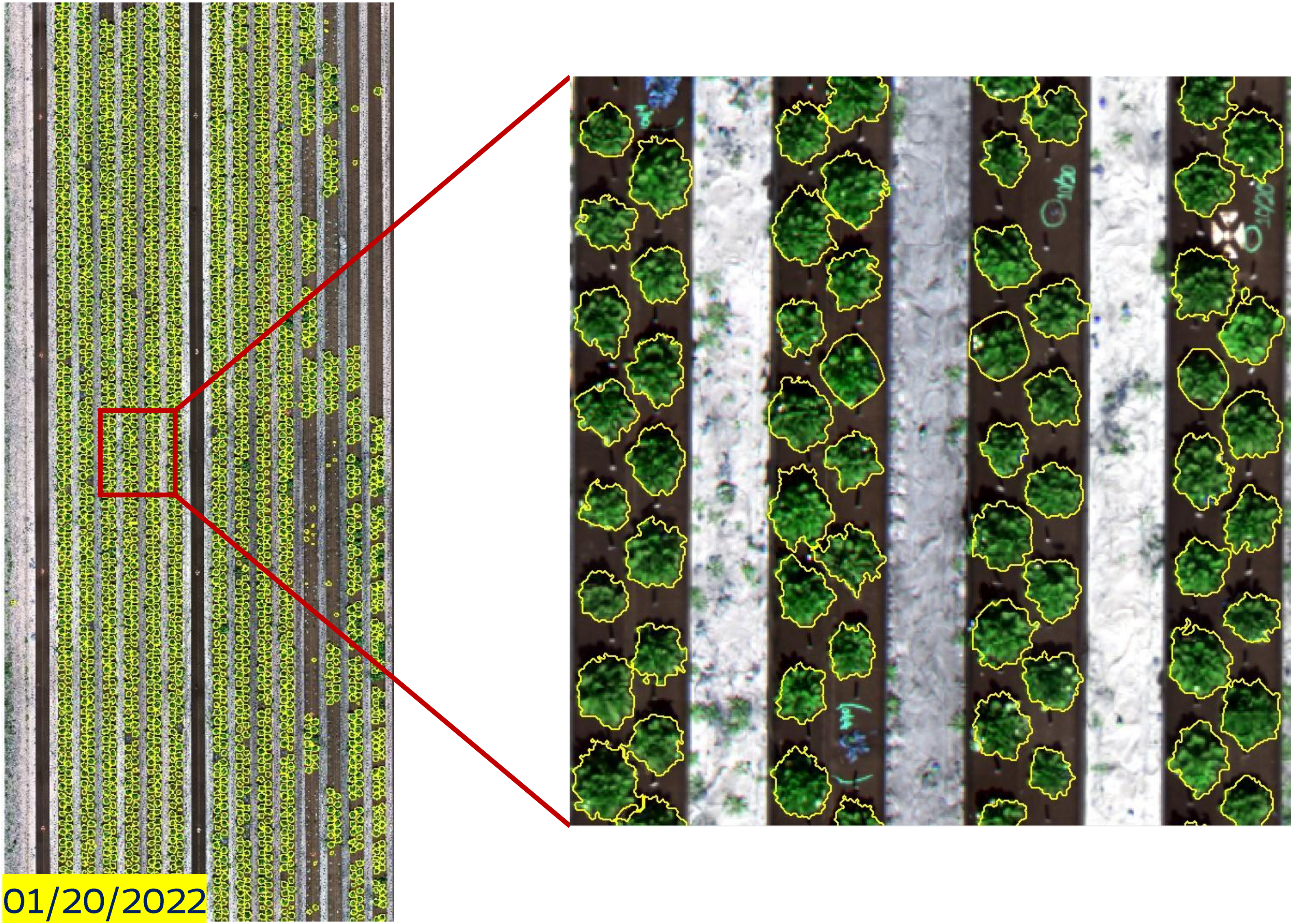


UF IFAS
UNIVERSITY of FLORIDA

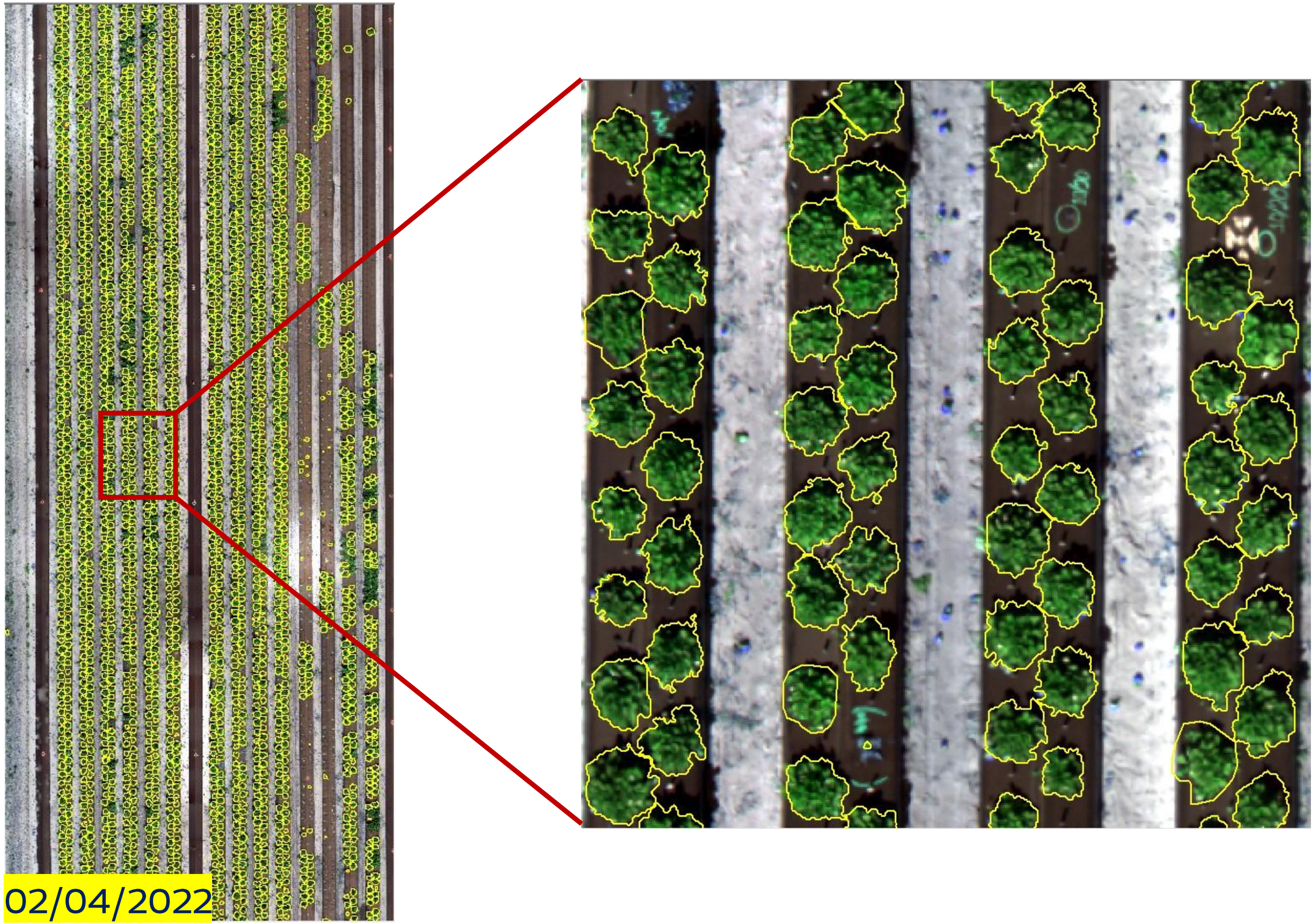
3). Strawberry Canopy Delineation using Deep Learning



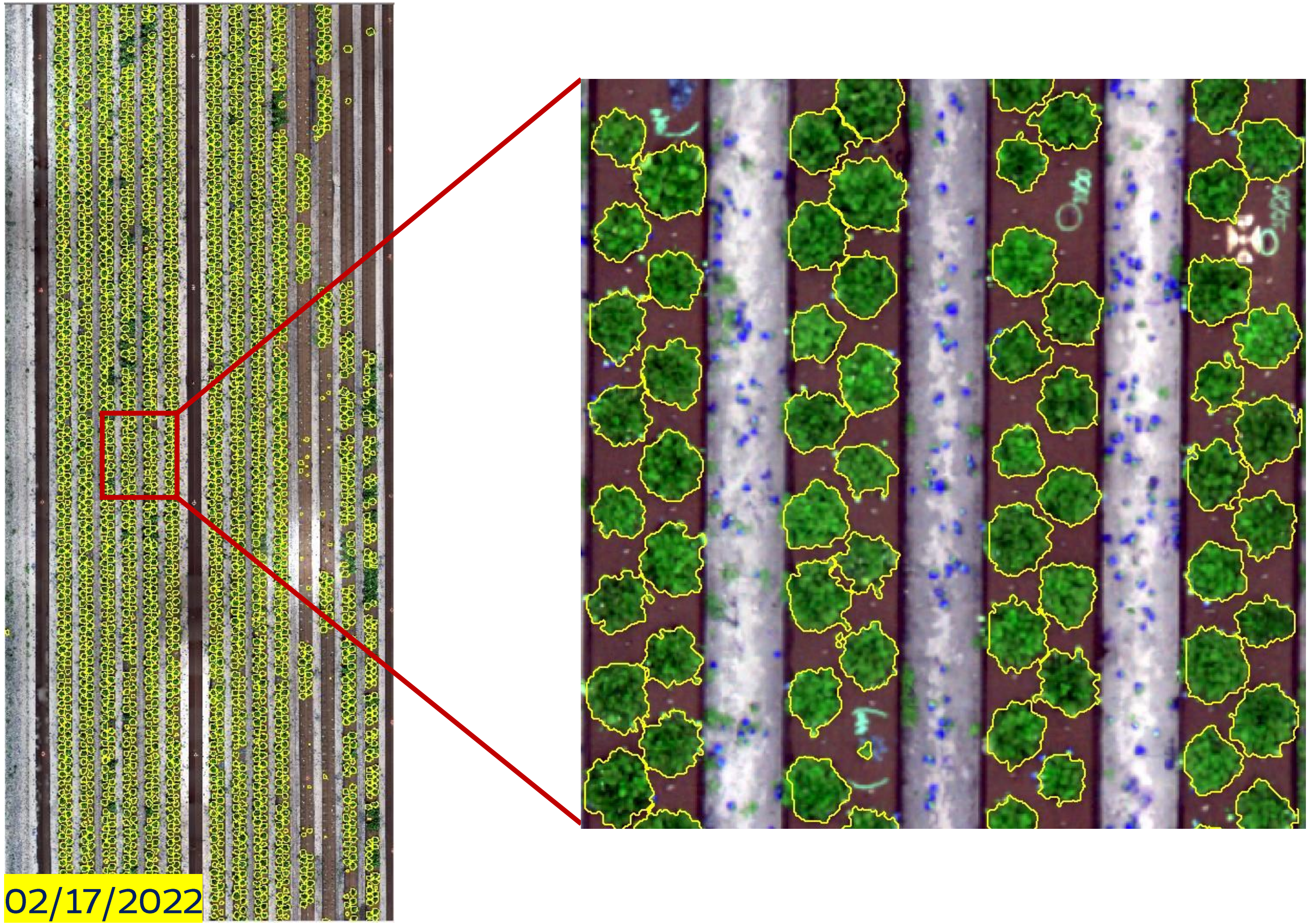
3). Strawberry Canopy Delineation using Deep Learning



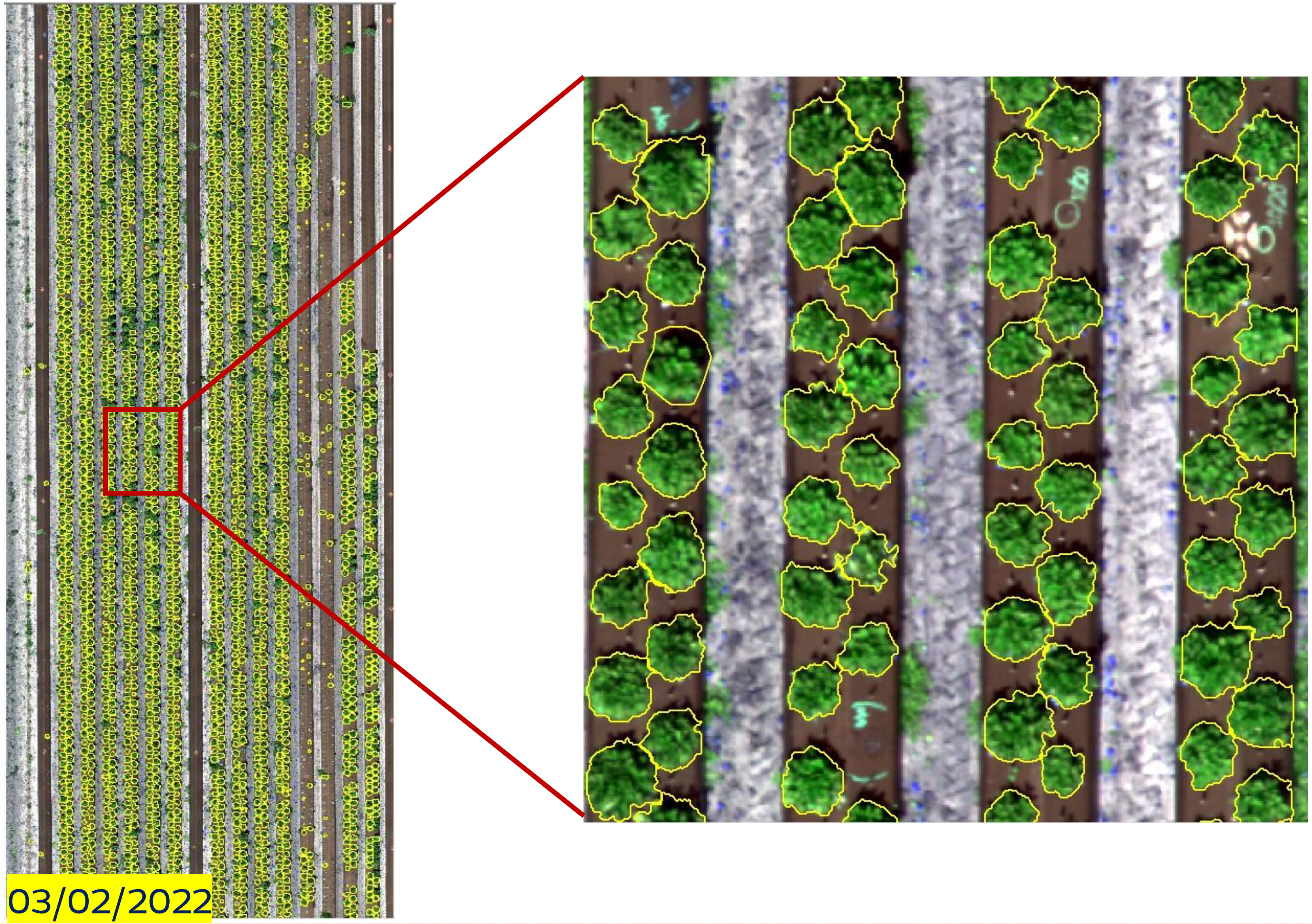
3). Strawberry Canopy Delineation using Deep Learning



3). Strawberry Canopy Delineation using Deep Learning



3). Strawberry Canopy Delineation using Deep Learning



4). Strawberry Canopy Delineation using Deep Learning

Table 2. Performance of various deep learning models

Method	mAP	AP_{50}	AP_{75}	AR
Mask R-CNN	0.709	0.979	0.889	0.745
YOLACT	0.682	0.917	0.758	0.652
SOLOv2	0.716	0.969	0.896	0.719
Mask2Former	0.689	0.907	0.841	0.691

Objective 2: Estimation of Strawberry Canopy Geometric Parameters from UAV Multispectral Imagery

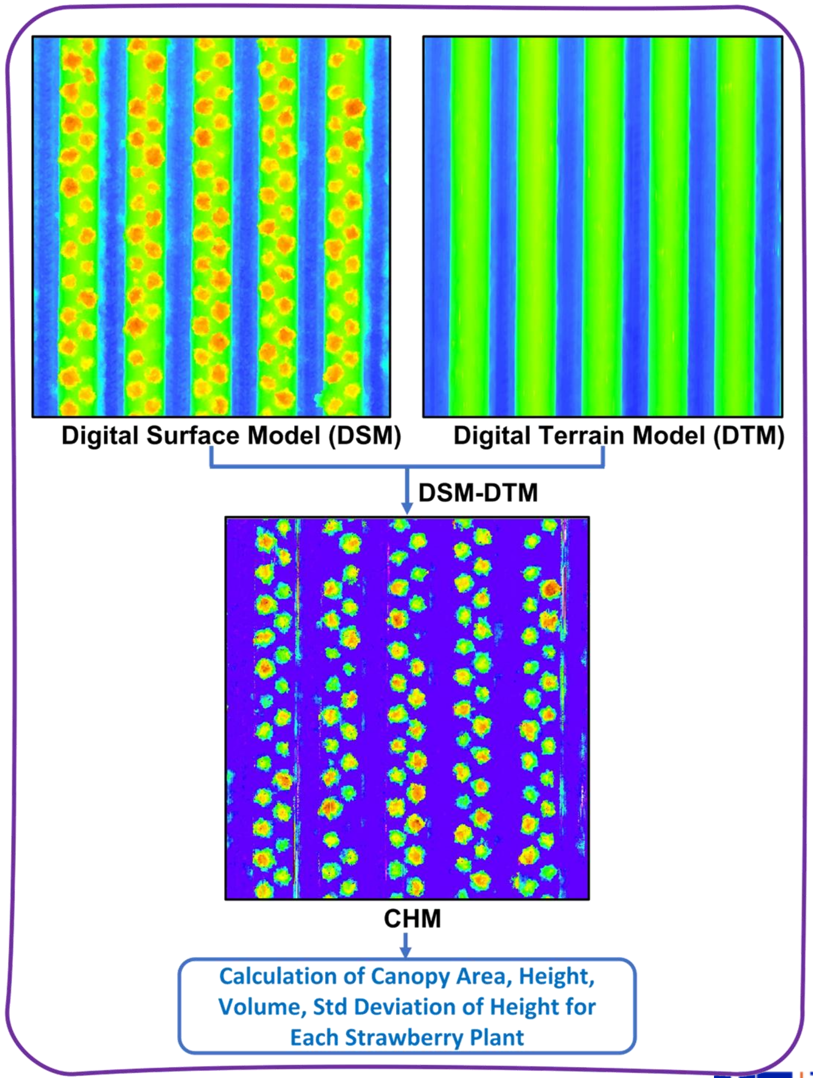
1). Strawberry Canopy Height Model Generation

The canopy height model (CSM) model can be obtained by subtracting the Digital Terrain Model (DTM) from the DSM, which was calculated as:

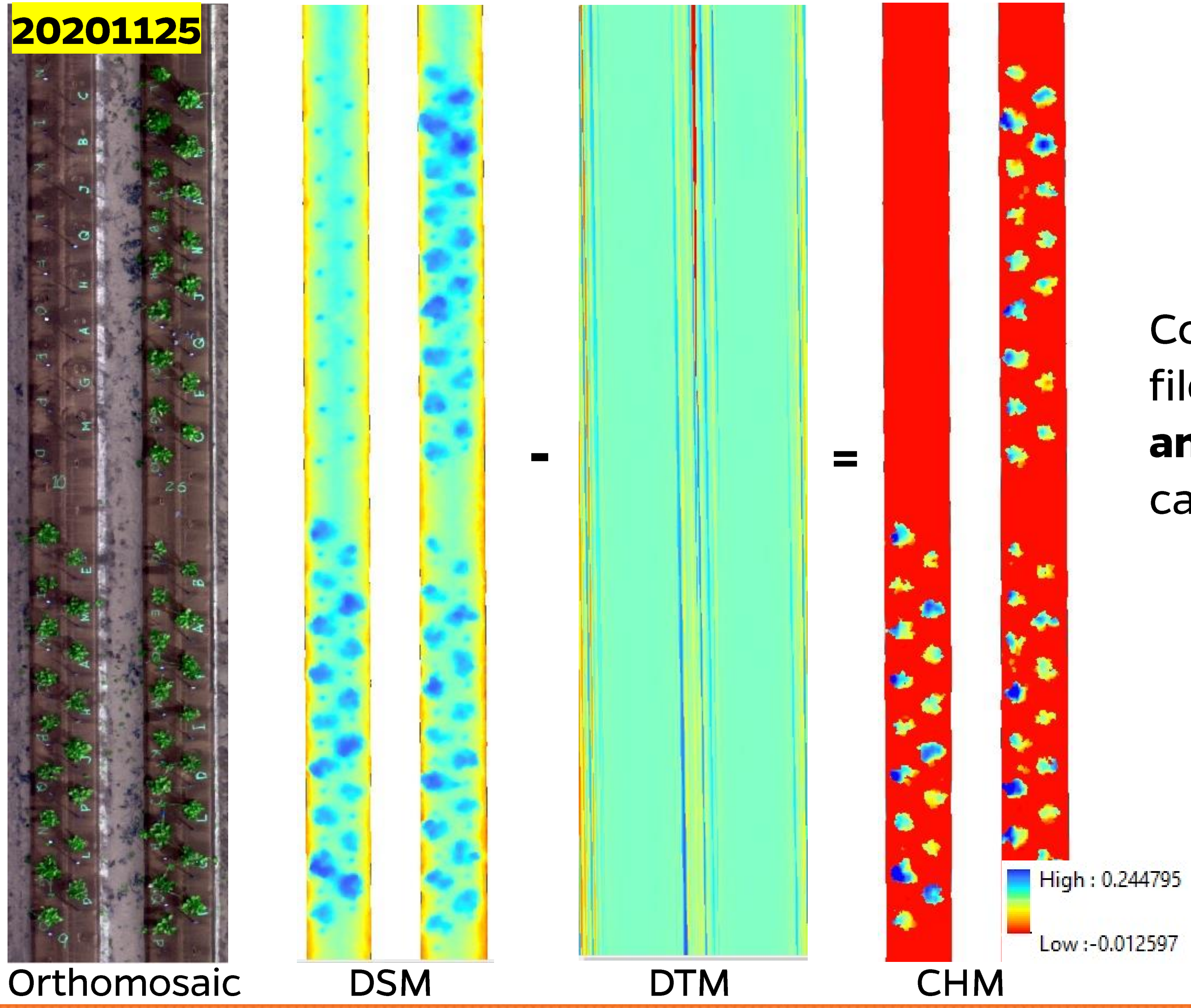
$$CSM = DSM - DTM$$

The DEM represents the elevation of the tops of objects in the image and is generated from the 3D point clouds in the SfM analysis. The DTM represents the elevation model that represents **ground surface elevation** from which the strawberry canopies are removed. **The vegetation-free DTM is interpolated using the elevations of the areas without vegetation cover by a local polynomial interpolation technique.**

Strawberry Canopy Height Modeling (CHM)



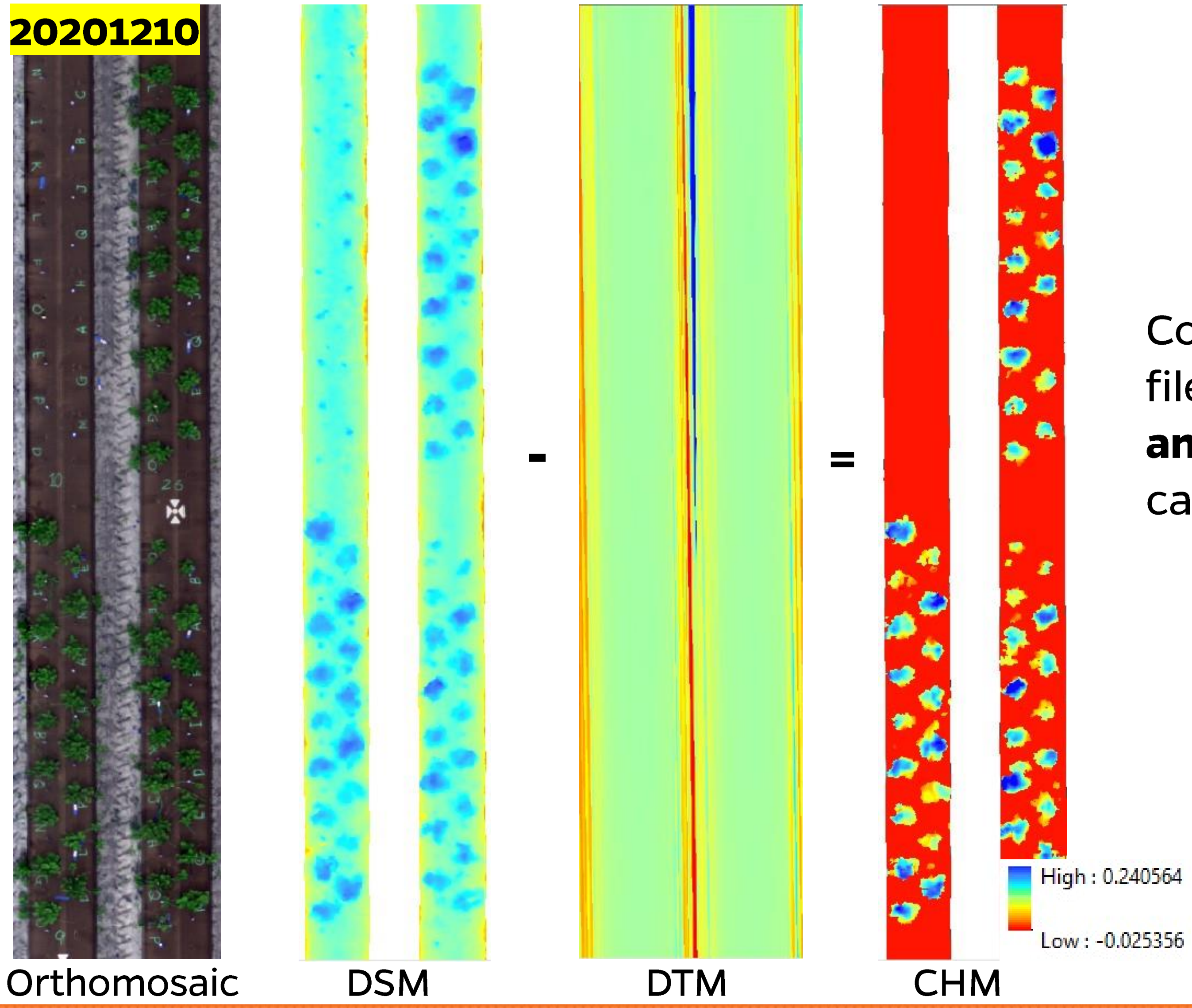
2). Estimation of Strawberry Canopy Geometric Parameters



Combined with the strawberry canopy boundary file, geometric parameters (**area, volume, height, and standard deviation of canopy height**) were calculated for each plant from the CSM model.

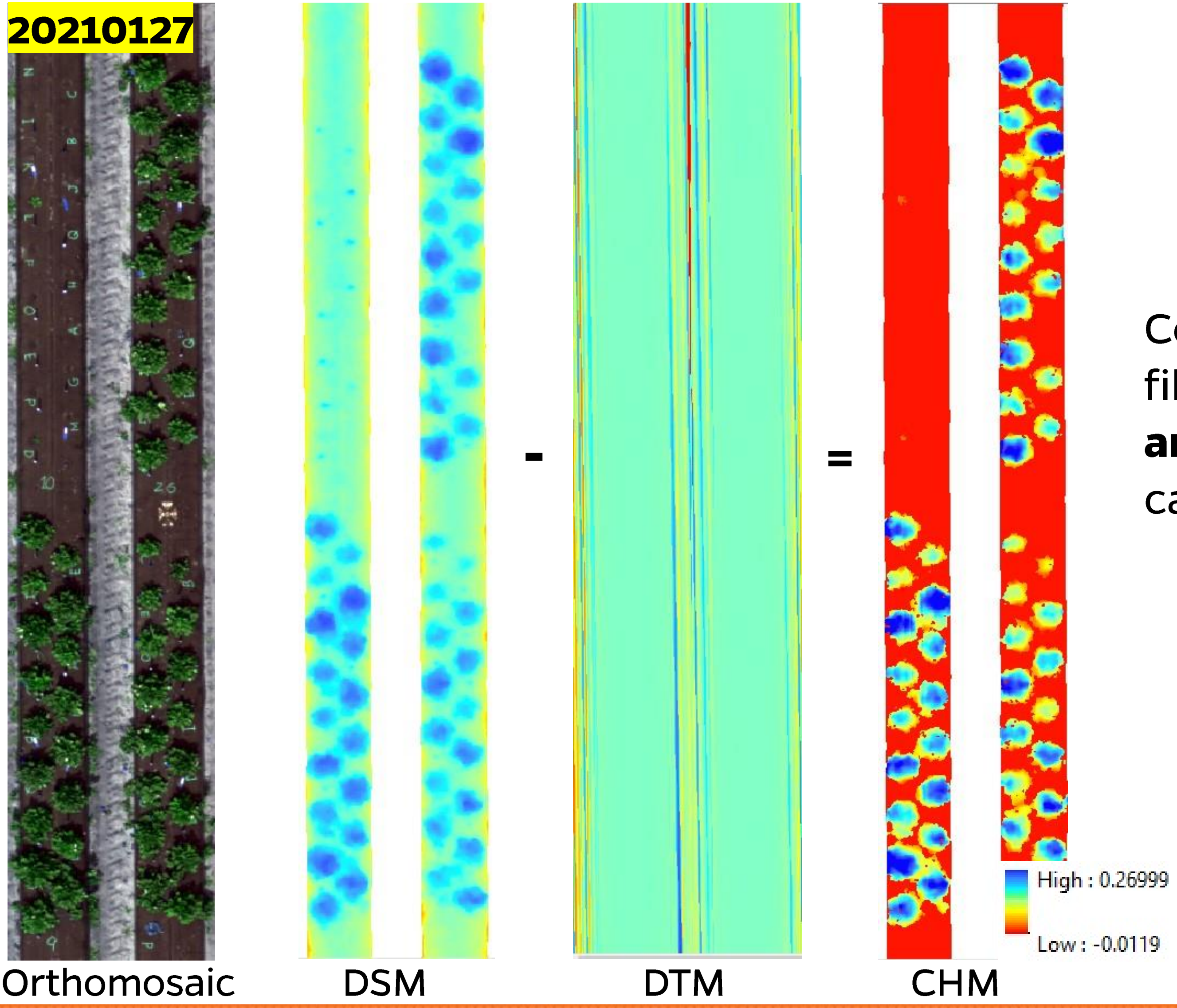
UF IFAS UNIVERSITY of FLORIDA

2). Estimation of Strawberry Canopy Geometric Parameters



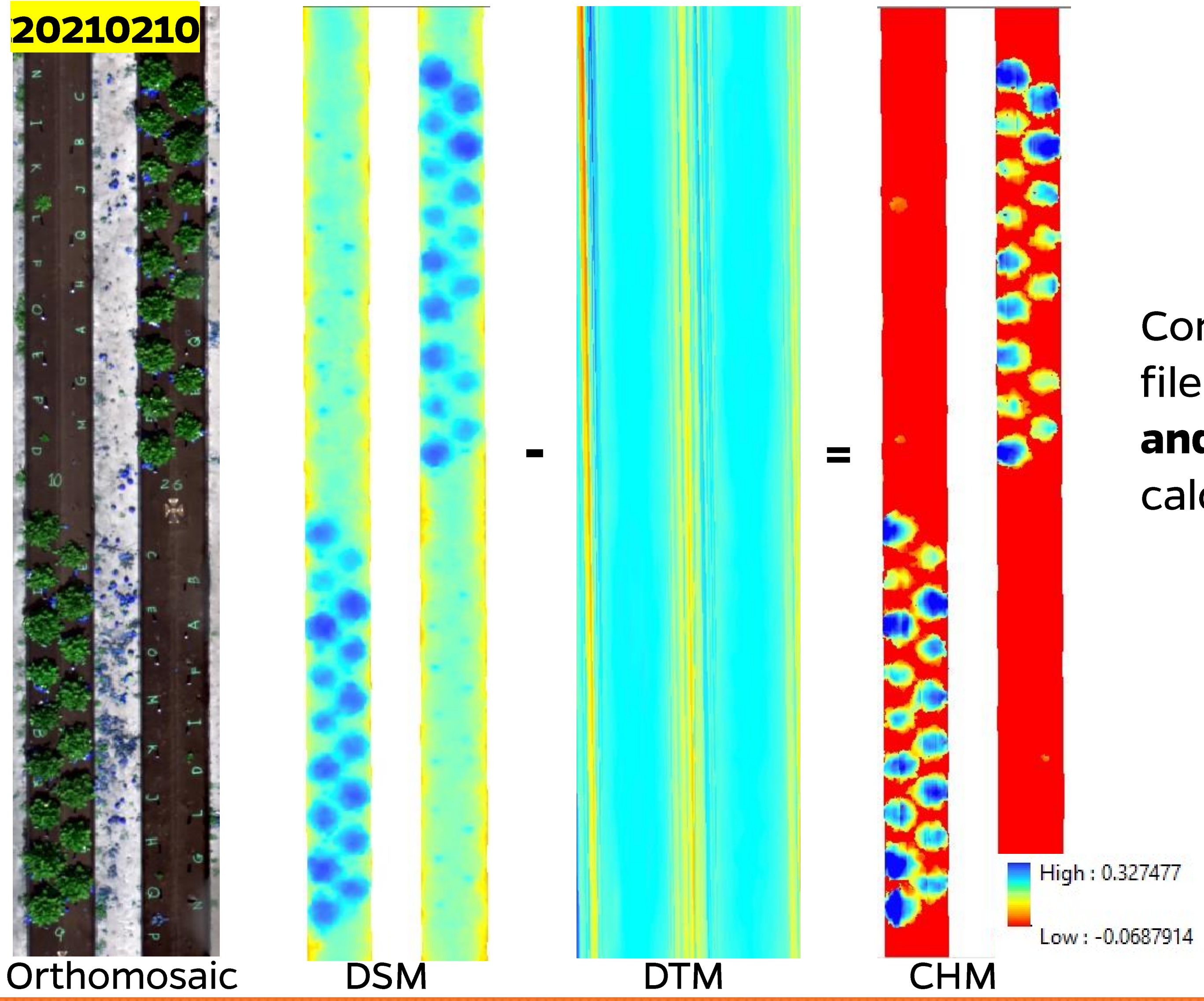
Combined with the strawberry canopy boundary file, geometric parameters (**area, volume, height, and standard deviation of canopy height**) were calculated for each plant from the CSM model.

2). Estimation of Strawberry Canopy Geometric Parameters



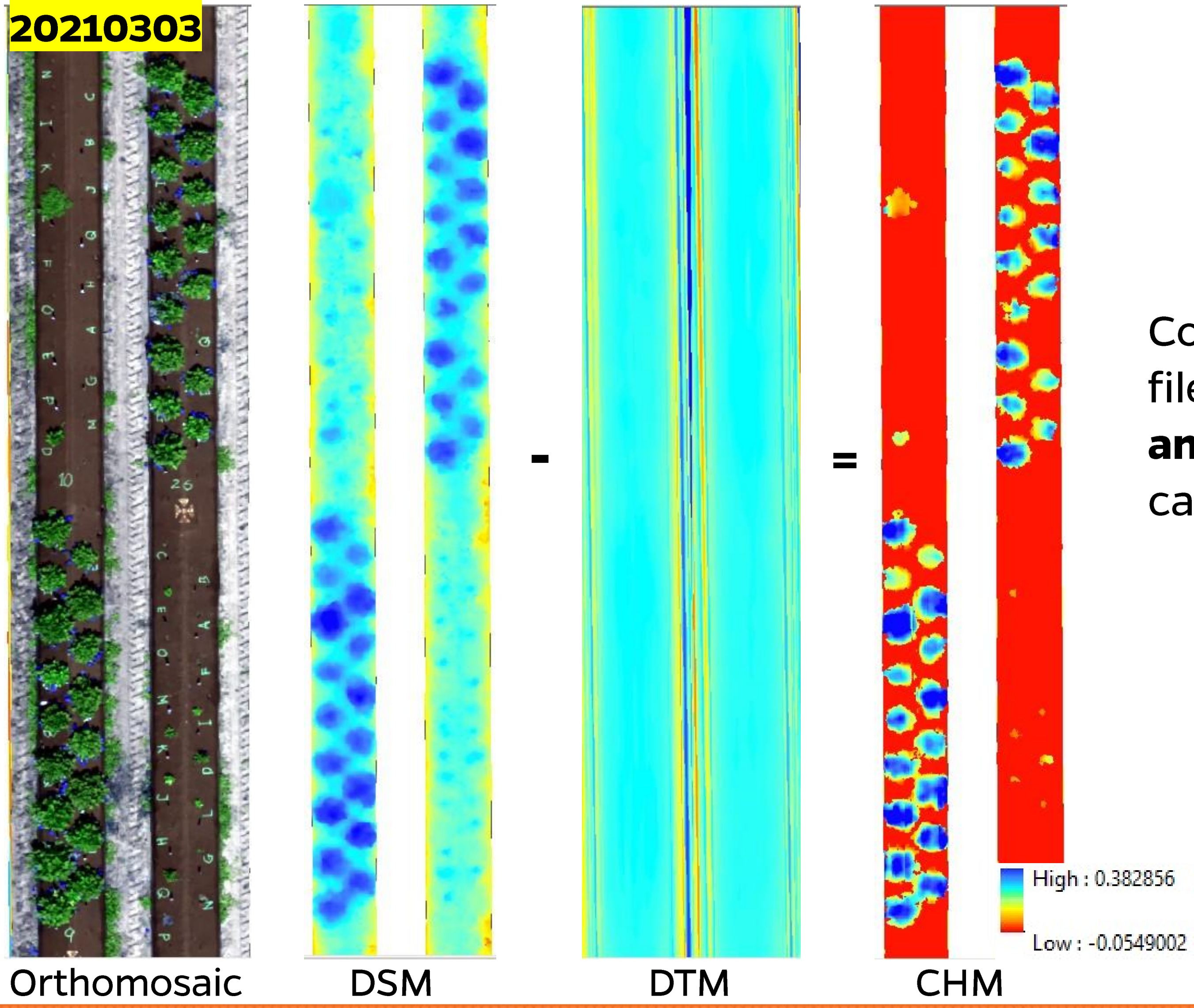
Combined with the strawberry canopy boundary file, geometric parameters (**area, volume, height, and standard deviation of canopy height**) were calculated for each plant from the CSM model.

2). Estimation of Strawberry Canopy Geometric Parameters



Combined with the strawberry canopy boundary file, geometric parameters (**area, volume, height, and standard deviation of canopy height**) were calculated for each plant from the CSM model.

2). Estimation of Strawberry Canopy Geometric Parameters



Combined with the strawberry canopy boundary file, geometric parameters (**area, volume, height, and standard deviation of canopy height**) were calculated for each plant from the CSM model.

3). Accuracy of Strawberry Canopy Geometric Parameters

Table 3. Descriptive summary of image-derived and in-situ dry biomass and height variables

Descriptive Statistics	Image-derived Variables				In-situ	
	Canopy area (m^2)	Height (m)	std Devidation of Height (m)	Volume (cm^3)	Dry biomass (g)	Height (m)
range	0.339	0.281	0.218	235.410	121.600	0.32
min	0.010	0.068	0.014	24.092	1.500	0.08
max	0.349	0.349	0.232	259.501	123.100	0.40
mean	0.104	0.176	0.047	73.784	28.904	0.19

The average height and dry biomass of the strawberry plant is about **18 cm and 29 g.**

UF IFAS
UNIVERSITY of FLORIDA

3). Accuracy of Strawberry Canopy Geometric Parameters

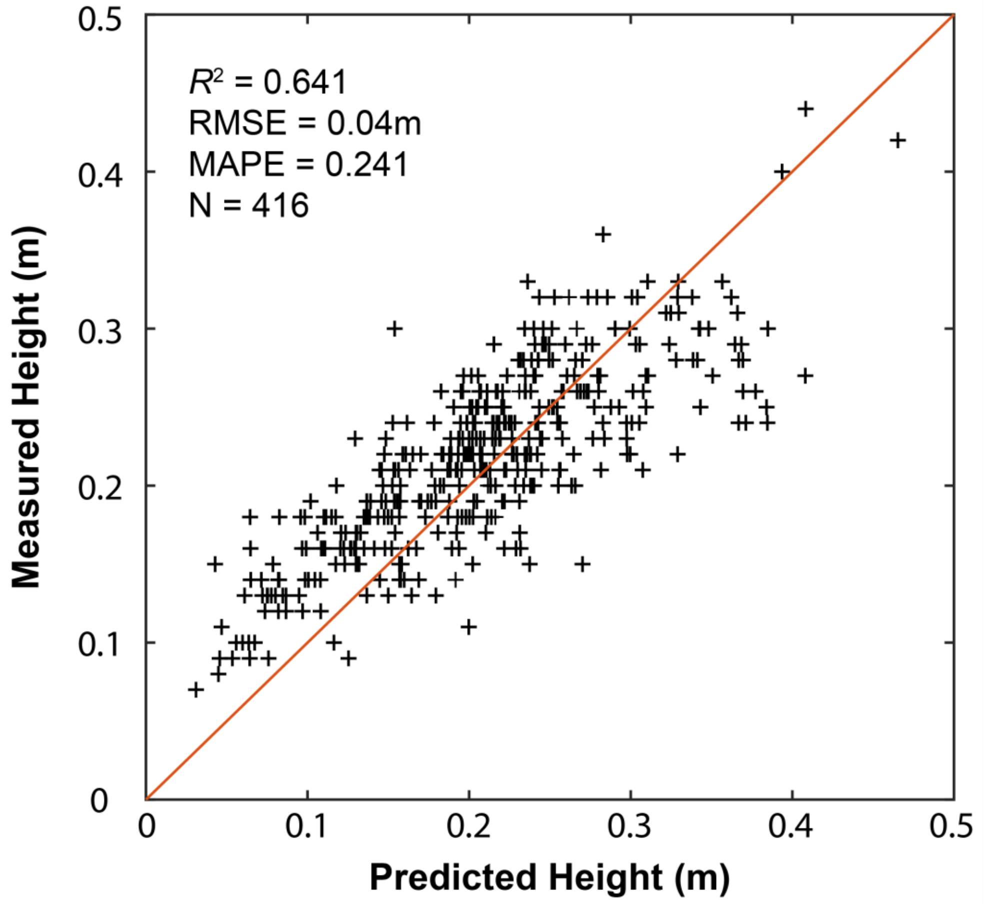


Figure 2. Scatter plots of the CHM-estimated plant height compared with the in-situ plant height data.

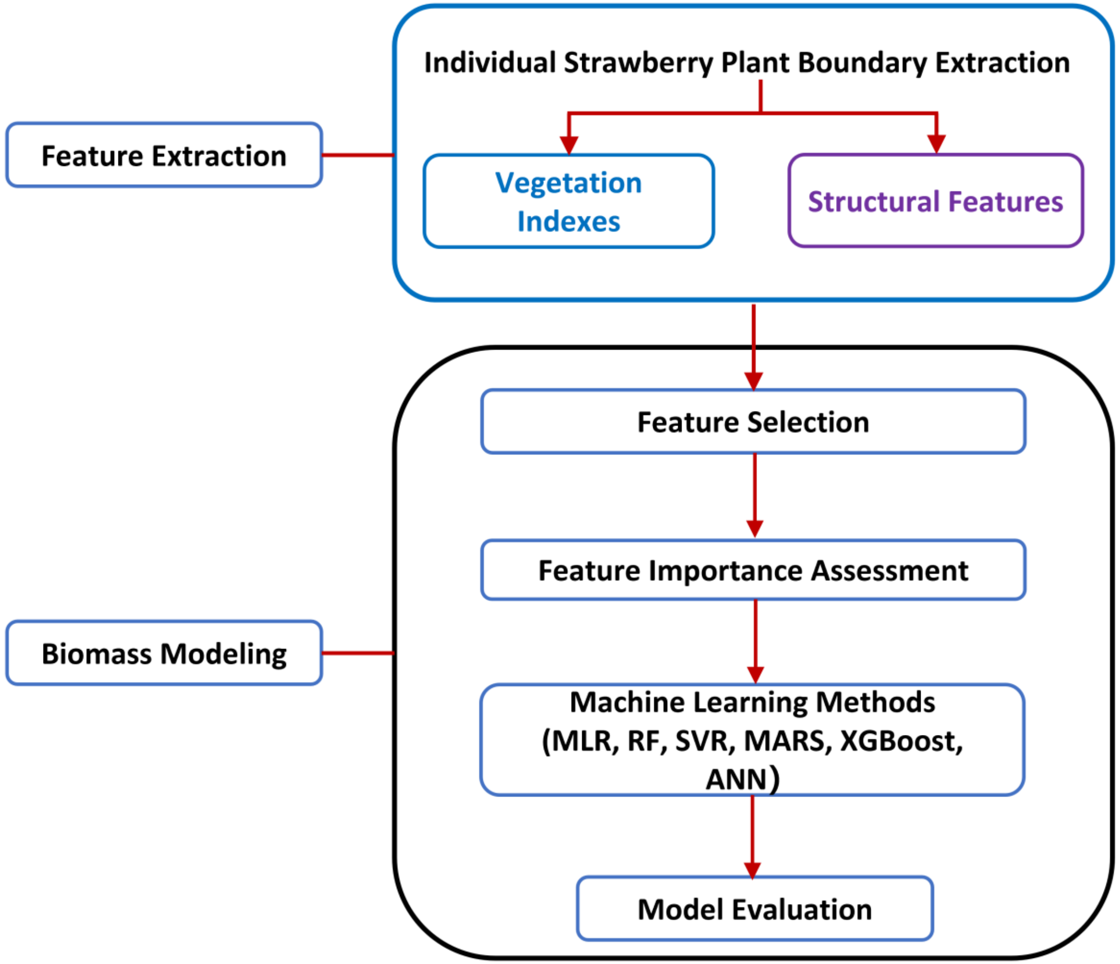
The CHM-estimated canopy height values were compared to the in-situ height data, as shown in Figure 2. The determination coefficient (R^2) between the image-derived heights using CHM and in-situ measured data was **0.641**, and the corresponding RMSE was about **4 cm**. The error source may come from the generated DSM and DTM.

Objective 3: Strawberry Dry Biomass Modeling Using Multiple Machine Learning Methods

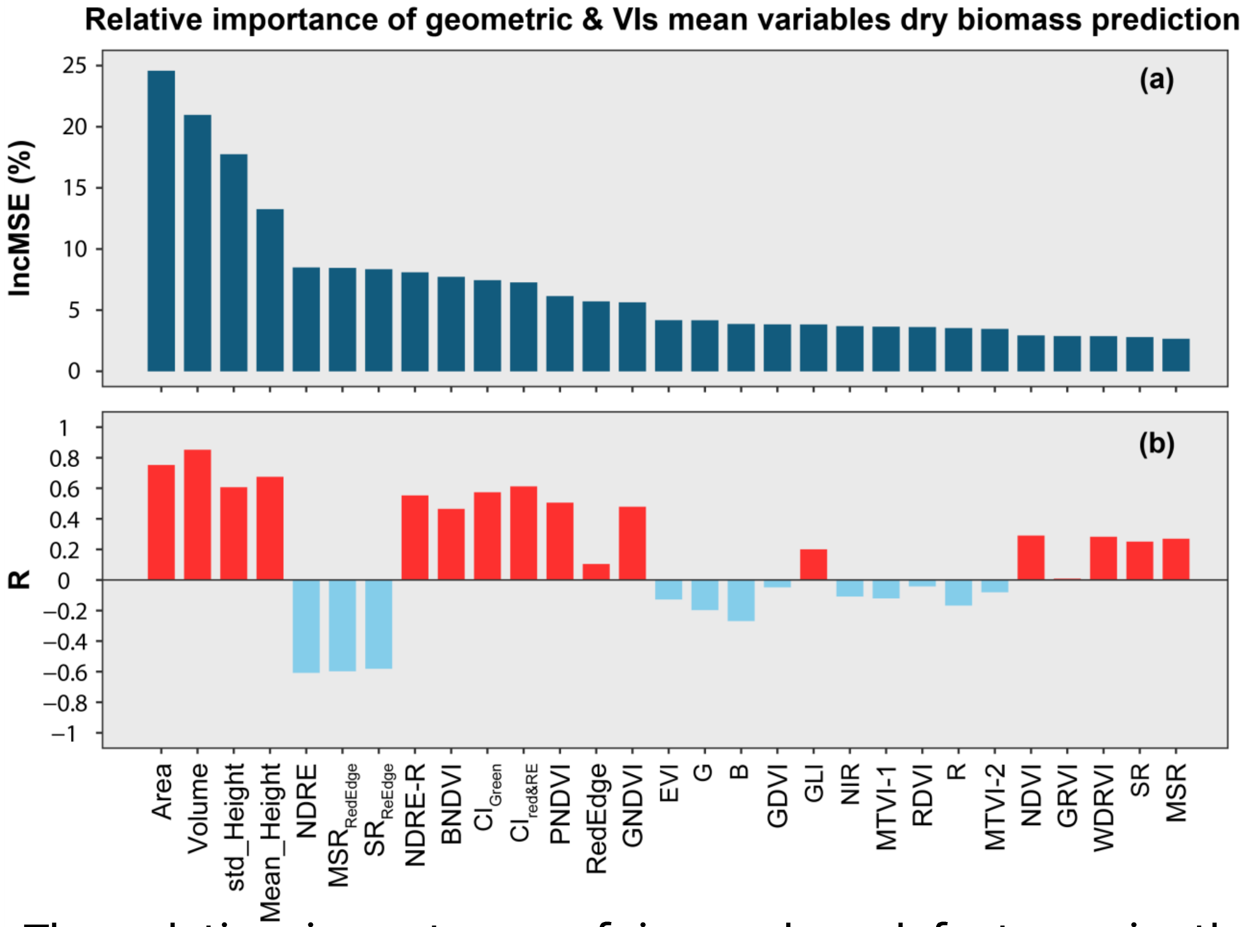
1). Strawberry Dry Biomass Modeling

Six regression models were used for strawberry dry biomass prediction: 1) Multiple linear regression (**MLR**), 2) Random forest (**RF**), 3) Support vector machine (**SVM**), 4) Multivariate adaptive regression spline (**MARS**), 5) XGBoost (**Extreme Gradient Boosting**), and 6) Artificial neural networks (**ANN**).

In addition to the **surface reflectance values of the five spectral bands**, we selected **20 VIs** that are commonly used to estimate canopy biophysical parameters. We generated a total of 29 dependent variables **(25 spectral features and 4 canopy structural variables**) for the biomass modeling.



2). Influential Features of Strawberry Biomass Modeling



Red-Edge related VIs including the normalized difference red-edge Index (**NDRE**), simple ratio vegetation index red-edge (**$SR_{RedEdge}$**), modified simple ratio red-edge (**$MSR_{RedEdge}$**) and chlorophyll index red and red-edge (**$CI_{red\&RE}$**) were the most influential VIs for biomass modeling.

The relative importance of image-based features in the prediction of strawberry biomass: a) Importance of geometric parameters & VIs mean variables for modeling dry biomass; b) Correlation coefficients of geometric parameters & VIs mean variables with dry biomass;

3). Predictive Performance of Six Machine Learning Methods

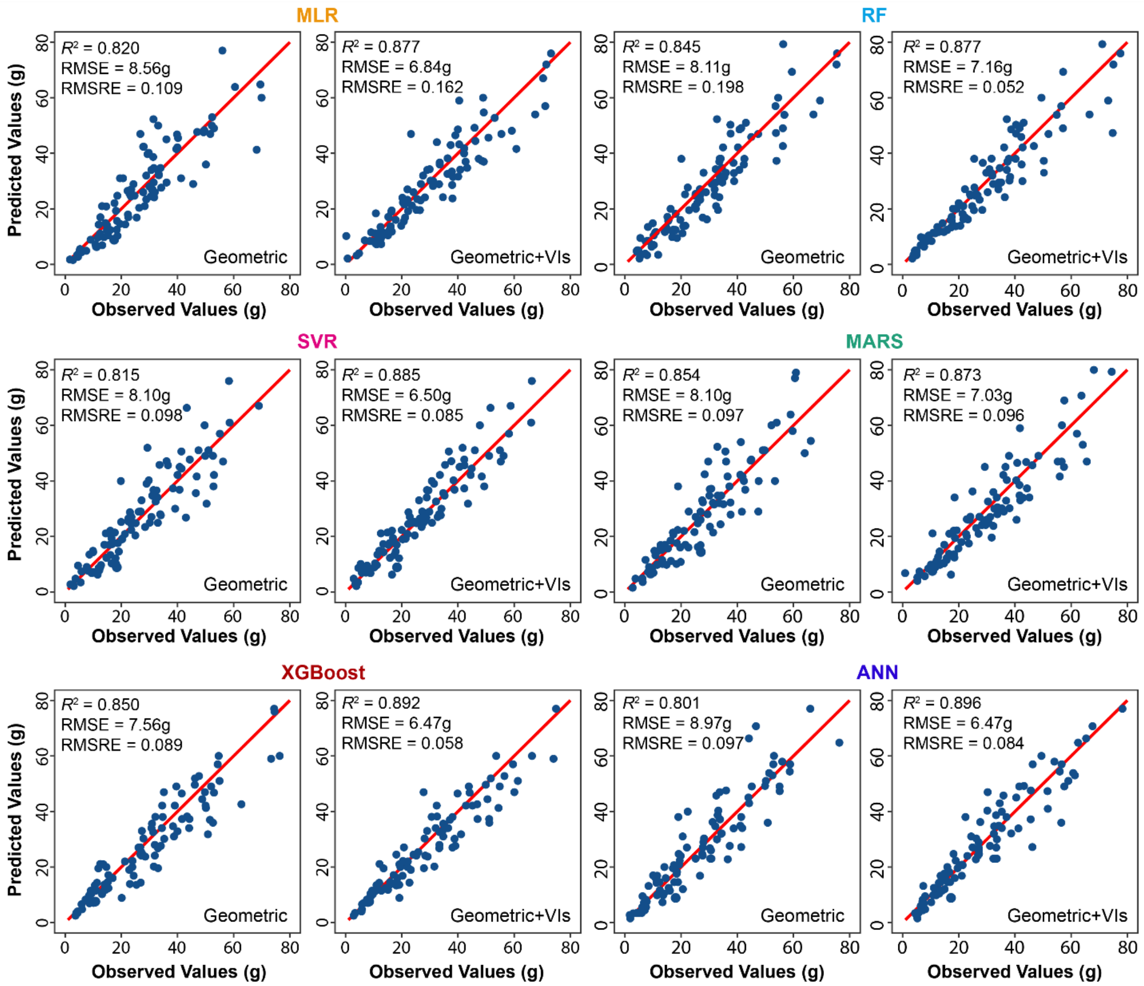
Table 4. Average five-fold CV results for the dry biomass models using the 2020-2021 dataset

Model	R^2			RMSE (g)			MAE (g)		
	Geometric	Geometric & VI_{mean}	Geometric & VI_{sum}	Geometric	Geometric & VI_{mean}	Geometric & VI_{sum}	Geometric	Geometric & VI_{mean}	Geometric & VI_{sum}
MLR	0.78	0.80	0.85	9.37	8.85	7.54	6.66	6.40	5.33
RF	0.80	0.81	0.85	8.89	8.72	7.73	6.30	5.95	5.37
SVM	0.78	0.80	0.85	9.58	8.92	7.67	6.58	6.09	5.26
XGBoost	0.77	0.82	0.83	9.41	8.67	8.09	6.70	5.85	5.47
MARS	0.79	0.81	0.86	9.27	8.80	7.35	6.53	6.14	5.25
ANN	0.89	0.90	0.93	8.98	8.61	7.16	6.29	5.93	5.06

The ANN had the highest accuracy in a five-fold cross validation with ***R*² of 0.89~0.93**, *RMSE* of **7.16~8.98 g**, and ***MAE* of 5.06~6.29 g**. As for the other five models, the addition of VIs increased the ***R*² from 0.77~0.80 to 0.83~0.86**, and reduced ***RMSE* from 8.89~9.58 to 7.35~8.09** g, and MSE from **6.30~6.70 to 5.25~5.47** g, respectively.

UF IFAS UNIVERSITY *of* FLORIDA

4). Predictive Performance of Six Machine Learning Methods



Scatterplots of in-situ measured plant dry biomass vs predicted values using 6 prediction models: MLR, RF, SVR, MARS, XGBoost and ANN.

UF IFAS UNIVERSITY of FLORIDA

Conclusion

In this study, we tried to extract strawberry canopy structural parameters and dry biomass from UAV multispectral imagery. We used computer vision and deep learning methods to extract canopy boundaries and estimate canopy structure parameters from the canopy height model (CHM). Finally, using the canopy structure parameters and vegetation indexes (VIs), we applied six regression models to predict the dry biomass with R^2 of about 0.8-0.9.

1. Zheng, C., Abd-Elrahman, A., & Whitaker, V. (2021). Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming. *Remote Sensing*, *13*(3), 531.
2. Zheng, C., Abd-Elrahman, A., Whitaker, V., & Dalid, C. (2022). Prediction of Strawberry Dry Biomass from UAV Multispectral Imagery Using Multiple Machine Learning Methods. *Remote Sensing*, *14*(18), 4511.
3. Zheng, C., Abd-Elrahman, A., Whitaker, V. M., & Dalid, C. (2022). Deep Learning for Strawberry Canopy Delineation and Biomass Prediction from High-Resolution Images. *Plant Phenomics*, *2022*.