

Beyond Segmentation: Outlier-Aware and Viewpoint-Aware Crop Water Stress Index Estimation in Dense Tomato Canopies

Laura Mas i Serra,^{0009-0008-3071-8685,a,1} Brian Kurzeja Jr.,^{a,1} Janaki Panneerselvam,^{0009-0000-1478-1141,b,1} Samuel Christopherson,^{0009-0009-8707-5454,a} Jacob Hampton,^{0000-0002-5904-8573,a} Eric Kor,^{0009-0001-8460-1043,a} Matthew Hatcher,^{0009-0009-6640-9523,a} Minas Pantelidakis,^{0000-0001-8908-0833,c} Konstantinos Mykoniatis,^b Orestis P. Panagopoulos,^d Shawn Ashkan,^e Athanasios Aris Panagopoulos^{0000-0003-1953-0100,a,2}

^aDepartment of Computer Science, California State University Fresno, Fresno, CA, USA

^bIndustrial and Systems Engineering Department, Auburn University, Auburn, AL, USA

^cIndustrial Engineering and Management Systems, University of Central Florida, Orlando, FL, USA

^dManagement Information Systems Department, California State University, Stanislaus, Turlock, CA, USA

^eCenter for Irrigation Technology, California State University Fresno, Fresno, CA, USA

apanagopoulos@mail.fresnostate.edu

Abstract—Precision irrigation depends on accurate crop water stress estimation, often using the Crop Water Stress Index (CWSI). Reliable CWSI estimation requires precise measurement of sunlit-leaf canopy temperature. Prior work introduced the Crop Irrigation Web Agent (CIWA) for convolutional neural network (CNN)-based sunlit-leaf segmentation and CWSI estimation in pistachio trees, focusing on sparse canopies. We extend CIWA to dense tomato canopies, where occlusion, reduced visual contrast, thermal leakage, residual RGB–thermal misregistration, and viewpoint effects make estimation more challenging. We introduce TMSL21, an annotated RGB-thermal dataset for tomato sunlit-leaf segmentation, compare CNN and Vision Transformer architectures, and propose a CWSI pipeline with Gaussian Mixture Model (GMM)-based filtering for outlier suppression. We also analyze imaging viewpoint effects, derive a tomato-specific CWSI baseline, and update CIWA for tomato imagery. Results show that reliable CWSI estimation depends on segmentation quality, thermal noise structure, and imaging geometry. Overall, the study generalizes CIWA toward a robust framework for dense canopies.

I. INTRODUCTION

OVER the next three decades, the global population is projected to increase by over 25%, resulting in a significant rise in demand for agricultural products [1], [2]. Most of these products depend on freshwater from surface and groundwater sources [3]. However, freshwater constitutes only a small fraction of total water resources and is becoming increasingly scarce in many regions as a result of climate change and overuse [4], [5], [6]. Effective freshwater management is therefore critical for a sustainable future, with precision irrigation playing a key role in optimizing water use [7].

A key component of many precision irrigation strategies is the Crop Water Stress Index (CWSI). This index quantifies

plant water stress based on environmental factors and sunlit-leaf canopy temperature (T_c), providing an estimate of the relative transpiration rate of a crop [8]. Accurate T_c measurements are therefore essential. A critical methodological consideration is whether to measure sunlit or shaded foliage. For irrigation purposes, T_c should be estimated from sunlit leaves, as these tend to exhibit water stress earlier and more strongly than non-sunlit leaves [9], [10]. However, direct collection of canopy temperature measurements at the individual-plant level is demanding for growers, requiring reliable instrumentation, standardized measurement procedures, and crop-specific baselines. In the context of thermal image analysis, the reliable isolation of sunlit leaves remains a major challenge. Manual identification of such regions is labor-intensive, difficult to scale, and prone to human error. Automated and accurate sunlit-leaf segmentation is therefore important not only for irrigation optimization, but also for broader applications such as microclimate modeling, physiological process analysis, and plant-disease monitoring [11], [12], [13].

Recent progress in this direction was established by the original Crop Irrigation Web Agent (CIWA) framework, which introduced supervised CNN-based sunlit-leaf segmentation for pistachio trees and integrated it into an expert system for CWSI estimation [14]. That work showed that segmentation precision is particularly important for downstream physiological estimation, since false positives may introduce shaded leaves, fruit, soil, sky, or other artifacts into canopy-temperature calculation. At the same time, the previous study focused primarily on pistachio imagery, where canopy structure is relatively sparse and image acquisition is typically performed from below. As already noted in that work, extending the framework to other crops such as tomato raises new challenges because smaller crops are commonly imaged from above, leading to different visible distributions of sunlit

¹ These authors contributed equally to this work.

² Corresponding author.

leaves and different error modes in downstream temperature estimation [14].

Tomato canopies provide a particularly challenging and revealing testbed exhibiting dense foliage, frequent occlusion, reduced visual contrast between sunlit and shaded leaves, and stronger exposure to background elements such as soil, in contrast with tree-based canopies. These factors make pixel-level segmentation more difficult and can distort downstream thermal analysis. This observation points to a broader limitation of existing approaches: segmentation is often treated as the end goal, rather than as one stage in a larger estimation pipeline whose final objective is reliable CWSI measurement. In practice, relatively small segmentation or residual RGB–thermal misregistration can induce heavy-tailed temperature distributions through thermal leakage, boundary contamination, or misclassified background pixels, while the imaging viewpoint can further reshape those distributions and bias crop-stress estimation.

In this paper, we extend CIWA from a segmentation-centered framework to a more robust end-to-end pipeline for dense tomato canopies. Our central claim is that reliable CWSI estimation is not determined solely by segmentation accuracy, but by the interaction between segmentation behavior, thermal noise structure, and imaging geometry. To study this problem, we introduce TMSL21, a co-registered RGB–thermal dataset for tomato sunlit-leaf segmentation; compare two convolutional neural networks (CNNs) and four Vision Transformer (ViT) architectures under dense-canopy conditions; and propose a Gaussian Mixture Model (GMM)-based likelihood filtering step to suppress influential thermal outliers before canopy-temperature aggregation. We further analyze the effect of top-view and side-view imaging on sunlit-leaf thermal distributions, derive a tomato-specific baseline for CWSI estimation, and extend the CIWA framework to support tomato imagery and pipeline-level evaluation.

Taken together, these contributions shift the focus from isolated segmentation performance to pipeline robustness for downstream physiological estimation. Our results indicate that segmentation-stage evaluation alone is insufficient for understanding downstream CWSI reliability, and that robust post-processing together with viewpoint-aware calibration is necessary for deployment in dense tomato canopies. In this sense, the present study should be understood not simply as a tomato-specific extension of CIWA, but as a step toward a more general and robust framework for thermal crop monitoring and crop water stress estimation.

The remainder of this paper is organized as follows. Section II reviews related work in crop water stress estimation and semantic segmentation for agricultural imagery. Section III introduces the TMSL21 dataset, including acquisition, annotation protocol, and dataset statistics. Section IV describes the proposed methodology, covering the extended CIWA pipeline, segmentation architectures, training configuration, evaluation metrics, and the proposed GMM-based outlier removal and viewpoint analysis. Section V presents the experimental results and discussion, including quantitative model comparisons,

thermal robustness analysis, and imaging viewpoint effects. Finally, Section VI concludes the study and outlines directions for future work.

II. RELATED WORK

A. Crop Water Stress Index (CWSI)

The Crop Water Stress Index (CWSI), originally introduced by Jackson et al. [8] and Idso [15], provides a quantitative measure of plant water status based on canopy temperature. The index is derived from the relationship between the canopy–air temperature differential ($T_c - T_a$) and atmospheric vapor pressure deficit (VPD). Under non-stressed conditions, transpiration induces evaporative cooling, resulting in lower canopy temperatures. In contrast, water-stressed plants exhibit reduced transpiration and consequently higher canopy temperatures.

To normalize this relationship across varying environmental conditions, two baselines are required: the non-water-stressed baseline (NWSB), representing the lower bound of ($T_c - T_a$) for well-watered plants, and the upper baseline (UL), corresponding to fully stressed conditions with minimal transpiration. The lower baseline is denoted as LL in this work and corresponds to the NWSB in prior literature. These baselines allow the calculation of a dimensionless index [15]:

$$CWSI = \frac{(T_c - T_a) - (T_c - T_a)_{LL}}{(T_c - T_a)_{UL} - (T_c - T_a)_{LL}} \quad (1)$$

where T_c denotes canopy temperature, T_a air temperature, $(T_c - T_a)_{LL}$ the lower baseline, and $(T_c - T_a)_{UL}$ the upper baseline. The resulting CWSI ranges from 0 (no water stress) to 1 (maximum water stress).

This formulation made CWSI largely environment-independent and established it as a standard tool for irrigation management [16]. Later work showed that reliable estimation depends critically on accurate T_c measurement, which can be improved by combining RGB imagery with thermal infrared data and by accounting for spatial variability within the canopy [9]. Gardner et al. further emphasized that infrared thermometry for irrigation should focus on sunlit leaves, as these respond to water stress earlier and more strongly than shaded foliage [10].

B. Sunlit-Leaf Segmentation

Accurate extraction of canopy temperature values (T_c) from sunlit leaves requires separating them from shaded foliage and background elements. Since manual annotation is impractical at scale, early computational approaches relied on color thresholding and Gaussian Mixture Model (GMM) refinement. For example, Wang et al. used a two-step process to isolate sunlit foliage from visible-spectrum imagery, but performance degraded under variable illumination and complex canopy conditions [17].

Advancements were made by Pantelidakis et al. [14], who introduced the first strictly supervised CNN-based framework for sunlit-leaf segmentation and integrated it into the Crop Irrigation Web Agent (CIWA) for pistachio-tree CWSI estimation. That study showed that CNN-based segmentation substantially

outperformed the earlier GMM-based approach, particularly in precision, which is crucial because false positives can contaminate canopy-temperature estimates with shaded leaves, sky, soil, or other non-sunlit regions [14]. However, that framework was evaluated primarily on pistachio trees, whose relatively sparse canopy structure differs substantially from dense tomato canopies.

C. Deep Learning Architectures for Agricultural Segmentation

Semantic segmentation has advanced significantly with deep neural networks. Convolutional neural networks (CNNs) remain widely used because of their strong ability to capture localized spatial features and fine-grained boundaries, which are important for identifying sunlit-leaf regions. In the context of CIWA, architectures such as FRRN-A and FC-DenseNet103 demonstrated strong performance on pistachio imagery [14]. More recently, Vision Transformer (ViT) architectures have emerged as competitive alternatives in segmentation tasks due to their ability to model long-range dependencies and global contextual information. However, their potential for sunlit-leaf segmentation in agricultural imagery remains underexplored. Beyond segmentation-focused applications, ML has shown strong performance in plant classification tasks. For example, a recent study reported 96.61% accuracy using KNN for plant species classification [18]. Deep learning has also shown strong predictive performance in environmental forecasting tasks, with Arzov et al. demonstrating that LSTM-based models outperforming ARIMA by incorporating pollution and meteorological data [19]. These findings further support the importance of environmental variables in robust CWSI estimation and the potential of data-driven approaches for crop-specific modeling.

D. Research Gap

Despite progress in sunlit-leaf segmentation, several limitations remain. First, prior supervised approaches have been validated mainly on tree crops with relatively sparse canopy structure, leaving open the question of generalization to dense crops such as tomato. Second, most existing approaches focus on segmentation quality as the primary endpoint, rather than on the robustness of downstream canopy-temperature and CWSI estimation. Third, prior pipelines rarely incorporate explicit filtering of extreme thermal values caused by segmentation errors, boundary contamination, or background leakage, even though such artifacts can bias downstream CWSI calculation. Finally, few studies have examined Vision Transformer (ViT) architectures for sunlit-leaf segmentation in agricultural imagery, particularly in dense canopy settings. The limitations mentioned above motivate the present study, which extends CIWA to dense tomato canopies through a comparison of CNN and ViT architectures, a revised annotated RGB-thermal tomato dataset, and an outlier-aware pipeline for robust downstream CWSI estimation.

III. TMSL21 DATASET

We release **TMSL21**, to the best of our knowledge the first annotated RGB-thermal dataset for tomato sunlit-leaf seman-

tic segmentation, publicly available at www.apanagopoulos.com/ciwa. The dataset was developed to support precise canopy temperature extraction for Crop Water Stress Index (CWSI) estimation using co-registered RGB and thermal imagery. TMSL21 contains 710 samples, each consisting of a synchronously captured and spatially aligned triplet:

- a 504×286 visible-spectrum RGB image of tomato plant,
- a 80×60 FLIR AX8 thermal image with embedded infrared channels and calibrated metadata, and
- a 504×286 binary mask labeling each pixel as either *sunlit leaf* or *non-sunlit leaf*.

RGB and thermal frames were acquired simultaneously using a FLIR AX8 camera, supporting pixel-level correspondence between modalities. The binary masks were generated through a semi-assisted annotation process described next.

A. Dataset Acquisition

The dataset was collected during the summer of 2021 in the San Joaquin Valley, California, USA, under clear-sky conditions around midday in order to minimize illumination variability. The FLIR AX8 camera was mounted for ground-based imaging at a fixed working distance, providing a native thermal ground sampling distance (GSD) of approximately 1.5 mm/pixel. Environmental metadata, including air temperature and relative humidity, were recorded using nearby weather stations. Camera-to-plant distance was manually measured and kept constant throughout acquisition, while crop emissivity and reflected background temperature were experimentally estimated following FLIR guidelines; all remaining parameters were kept at FLIR AX8 default settings [20]. To enable pixel-level alignment with RGB images, the native 80×60 thermal frames were upscaled to 504×286 using bilinear interpolation.

B. Dataset Annotation

Each RGB image was manually annotated by trained experts to produce a binary semantic mask labeling pixels as either *sunlit leaf* or *non-sunlit leaf*. To improve efficiency and annotation consistency, unsupervised clustering techniques were used as a preprocessing aid. In particular, an adapted version of the open-source annotation tool Pynovisao, previously used in [14] and originally introduced by Dos Santos et al. [21], was employed to support the identification of sunlit-leaf regions. The tool groups pixels with similar visual characteristics, helping annotators distinguish sunlit foliage from shaded leaves and background elements through a click-to-select workflow. The adapted version was extended including support for Python 3.12, improved ground-truth generation through the combination of SLIC-based superpixels with thermal information, and annotation guidelines to support consistent labeling. The latest version of the Pynovisao tool is publicly available at <https://github.com/bkur808/Sunlit-leaf-annotator>. Finally, all generated masks underwent quality assurance through an automated overlay procedure in which each mask was superimposed on its corresponding RGB image for visual verification.

TABLE I
SUMMARY STATISTICS OF THE TMSL21 DATASET

Property	Value
Total samples	710
RGB resolution	504×286 px
Thermal resolution (native)	80×60 px
Thermal resolution (upsampled)	504×286 px
Thermal GSD	~ 1.5 mm/pixel
Annotation type	Binary
Acquisition method	Ground-based imaging
Environmental conditions	Clear-sky
Capture time window	Approximately midday
Imaging device	FLIR AX8
Training samples	497 images (70%)
Validation samples	142 images (20%)
Test samples	71 images (10%)

C. Dataset Statistics and Samples

The main statistical properties of TMSL21, including sample distribution and imaging specifications, are summarized in Table I. Representative samples are shown in Figure 1, where (a) presents RGB imagery, (b) thermal imagery, and (c) the corresponding binary mask.

D. Dataset Validation and Correction

During post-processing and visual inspection of temperature-mapped overlays, we identified inconsistencies in a subset of samples in which duplicated visible-spectrum images had been paired with different thermal frames, resulting in invalid temperature-to-pixel correspondences. To ensure dataset integrity, all RGB–thermal pairings were re-verified. In total, 98 affected samples (approximately 12% of the original dataset) were excluded from the final release. All experiments reported in this study were conducted exclusively on the filtered and verified dataset. By combining high-quality imagery, verified RGB–thermal pairings, and environmental metadata, TMSL21 provides a robust foundation for research on CWSI estimation, sunlit-leaf temperature analysis, and precision irrigation.

IV. METHODOLOGY

A. Expanded CIWA Pipeline for Tomato Processing

Building upon the original CIWA framework [14], which was designed for tree-based crops using CNN-based segmentation, we introduce an expanded methodology tailored to dense-canopy crops such as tomato. The proposed pipeline extends CIWA in three directions. First, we compare multiple segmentation architectures, including convolutional neural networks and Vision Transformer-based models, to assess their suitability under dense foliage conditions. Second, we incorporate a probabilistic outlier-suppression step based on Gaussian mixture modeling (GMM) in order to reduce the impact of extreme thermal artifacts on canopy temperature estimation. Third, we examine the effect of imaging viewpoint

on the thermal distribution of predicted sunlit-leaf pixels and its downstream implications for CWSI estimation. The overall pipeline is illustrated in Figure 2.

B. Thermal–RGB Alignment for Temperature Extraction

The TMSL21 dataset provides co-registered RGB and thermal imagery, supporting pixel-level correspondence between visual features and temperature values. This alignment is critical for CWSI estimation, since false positives in sunlit-leaf segmentation may introduce non-sunlit regions into the canopy-temperature calculation.

Sunlit-leaf segmentation is first performed on RGB images, and the resulting masks are mapped onto aligned thermal frames to extract the corresponding temperature values. The canopy temperature is then computed as the mean temperature over the predicted sunlit-leaf region.

Although all RGB–thermal pairings were verified, residual local misalignment was frequently observed, likely due to sensor geometry, thermal upsampling, parallax, and canopy depth variation. As a result, visually correct sunlit-leaf masks may overlap neighboring thermal pixels from soil, shaded foliage, or canopy boundaries, contributing to heavy-tailed temperature distributions.

To improve robustness against these effects, a GMM-based likelihood filtering step is subsequently applied to remove extreme thermal outliers before aggregation, thereby reducing the influence of segmentation errors, boundary artifacts, and misalignment-induced thermal artifacts.

C. Segmentation Architectures

The architectures evaluated in this study represent two main families of modern segmentation approaches: convolutional neural networks (CNNs) and Vision Transformers (ViTs). CNN-based models, including FRRN-A [22] and FC-DenseNet103 [23], were selected due to their demonstrated effectiveness in canopy segmentation, particularly in prior pistachio-based CIWA studies [14].

Vision Transformer-based models were included to assess the impact of global self-attention mechanisms on dense agricultural imagery. These include SegFormer [24] and Mask2Former [25], both of which have achieved strong performance on large-scale benchmarks such as Cityscapes and ADE20K [26], [27]. SegFormer employs a hierarchical transformer encoder with a lightweight decoder, while Mask2Former uses a Swin Transformer backbone with a mask-classification framework for pixel-level prediction.

In addition, we include foundation-model baselines to assess transferability to agricultural sunlit-leaf segmentation. The Segment Anything Model (SAM) [28] is a promptable segmentation model trained on large-scale heterogeneous datasets and is used here without task-specific fine-tuning to assess generalization to sunlit-leaf regions. DINOv3 [29] is included as a self-supervised vision transformer to evaluate the usefulness of learned visual representations in a low-data agricultural setting.

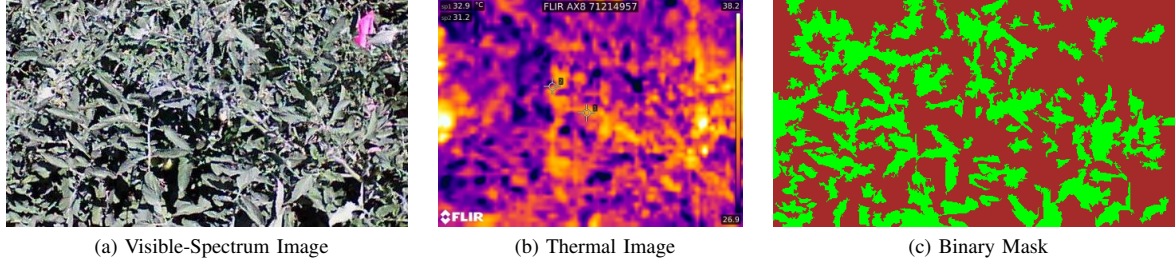


Fig. 1. Dataset Samples

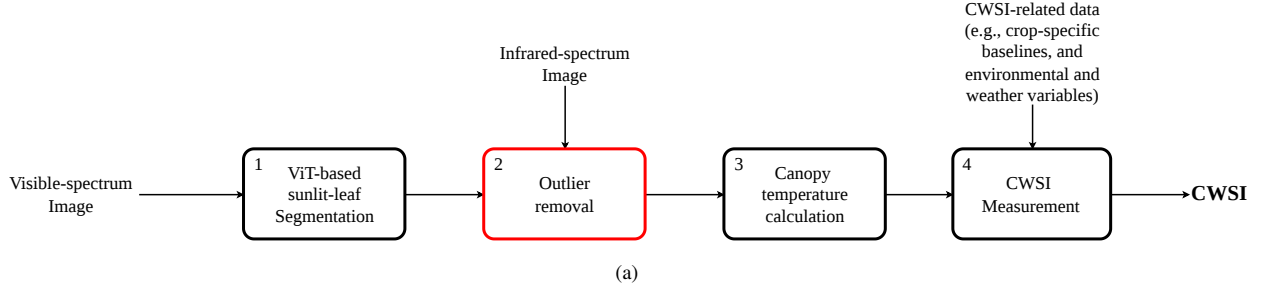


Fig. 2. Expanded CIWA pipeline for tomato processing and CWSI estimation (adapted from [14]).

The evaluation metrics used in this study are defined in Section IV-E, and corresponding experimental results and comparative analysis across all models are presented in Section V-A.

D. Training Configuration

Experiments were conducted using four NVIDIA RTX 6000 GPUs (49 GB VRAM each) with CUDA 12.8. The models were implemented in PyTorch and trained with data parallelism across GPUs for a maximum of 100 epochs, using a total batch size of 16 images (4 images per GPU). A fixed random seed (42) was used to improve reproducibility.

The TMSL21 dataset was randomly partitioned into training (497 images), validation (142 images), and test (71 images) subsets. All models were trained using a pixel-wise softmax cross-entropy loss between per-label activations (logits) and ground-truth labels. Optimization was performed using Root Mean Square Propagation (RMSProp) with a learning rate of 0.0001 and a decay factor of 0.995. The validation set was used to monitor generalization and implement early stopping in order to reduce the risk of overfitting [30]. Whenever validation loss improved, model weights were saved as a checkpoint. Otherwise, a patience-based early-stopping criterion was applied, terminating training when validation loss failed to improve for seven consecutive validation rounds. The checkpoint achieving the lowest validation loss was subsequently used for evaluation on the test set.

E. Segmentation Evaluation Metrics

The performance of each approach was evaluated using accuracy, precision, recall, F1 score, and mean Intersection-over-Union (IoU). Among these, precision is the priority metric,

as it quantifies the reliability of predicted sunlit-leaf pixels and directly influences the accuracy of canopy temperature (T_c) and Crop Water Stress Index (CWSI) estimation. IoU is not sufficient as a standalone metric in this context, since it assigns equal weight to false positives and false negatives, whereas CWSI computation is more sensitive to false positives that introduce non-sunlit pixels into T_c estimation, resulting in systematic underestimation of plant stress. Recall captures segmentation coverage but has reduced impact on bias, as missed sunlit regions primarily affect sample size rather than introducing significant temperature distortion. Accuracy and F1 score are reported for completeness:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Mean IoU} = \frac{1}{2} \left(\frac{TP}{TP + FP + FN} + \frac{TN}{TN + FP + FN} \right)$$

F. Handling High-Temperature Outliers in Sunlit Leaves

The thermal distribution of tomato sunlit leaves reported in Figure 3a exhibits a right-skewed tail, which can inflate canopy temperature T_c when computed as the average of predicted sunlit pixels. Consequently, this effect can propagate to the Crop Water Stress Index (CWSI) and lead to overestimation of plant stress. By contrast, the thermal distribution of pistachio

sunlit leaves, obtained using PTSL21 [14], in Figure 3c exhibits a left-skewed (cold) tail, highlighting crop-dependent differences in thermal artifact structure.

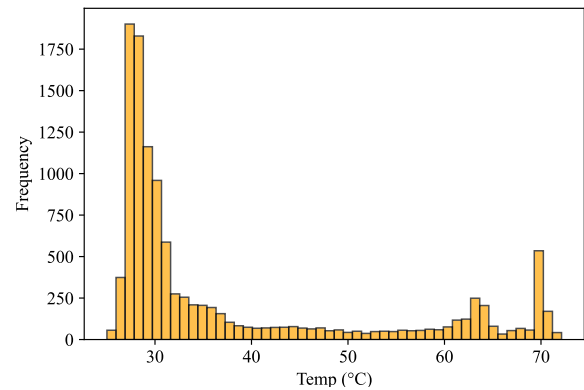
Spatial analysis of the corresponding overlays reveals that these extreme values are not uniformly distributed across the canopy. In tomato imagery, as shown in Figure 3b, high-temperature outliers (65–70 °C) are concentrated along canopy boundaries and in sparsely covered regions, forming irregular edge-aligned patterns. In pistachio imagery, pictured in Figure 3d, cold outliers (approaching -40°C) are localized near the upper canopy edges, where the foliage interfaces with the sky, appearing as fragmented bands along the canopy boundary.

These spatial patterns suggest that the dominant source of extreme values in both cases is not physiological variation alone, but a combination of segmentation error, boundary contamination, and residual RGB–thermal misregistration. In tomato images, the boundary-concentrated hot pixels are consistent with bare-soil contamination, where highly heated soil is misclassified as sunlit foliage. In pistachio images, the presence of extremely low temperatures near canopy tops indicates sky pixels leaking into the leaf mask. The structured, boundary-driven nature of these artifacts distinguishes them from true physiological signals, which are expected to occur within the canopy interior, and from sensor noise, which is typically more spatially diffuse.

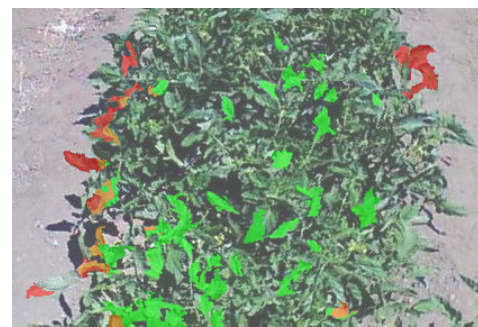
Before applying any outlier-removal strategy, it is important to distinguish between different sources of extreme values. First, elevated temperatures may reflect genuine physiological stress in sunlit leaves and therefore carry meaningful information for CWSI estimation. Removing such values indiscriminately could lead to underestimation of plant stress. Second, segmentation artifacts such as soil, branches, or fruit can introduce artificial temperature tails, particularly near image boundaries. Third, thermal sensor noise and measurement uncertainty, including reflective effects and emissivity mismatch [31], can also inflate values. These issues are especially important in field infrared measurements, where unaccounted background radiation may introduce systematic bias of 1–3°C [32]. Finally, residual RGB–thermal misregistration can also contribute to heavy-tailed temperature distributions by shifting predicted canopy-boundary pixels onto thermally distinct regions such as soil, sky, or shaded foliage. This underscores the need for selective outlier handling strategies that account for spatial context, rather than relying on global filtering approaches that may remove physiologically meaningful temperature extremes.

G. Implementation of a GMM-Based Method for Outlier Removal

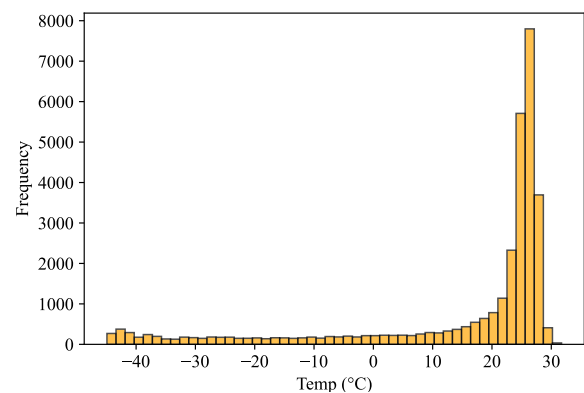
To address sensitivity to extreme values in the sunlit-leaf temperature distribution, we introduce a probabilistic outlier-removal strategy based on a Gaussian Mixture Model



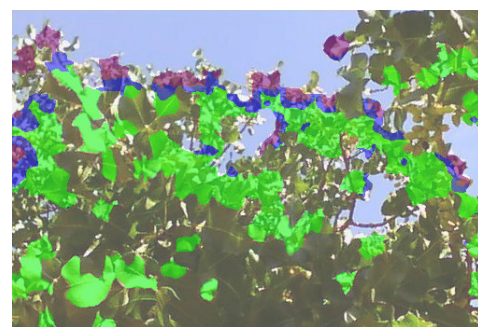
(a) Tomato sunlit T_c distribution



(b) Tomato thermal mapping

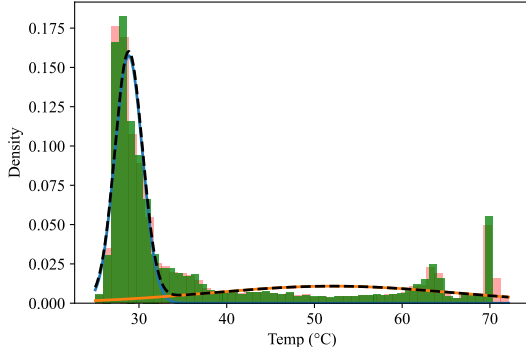


(c) Pistachio sunlit T_c distribution



(d) Pistachio thermal mapping

Fig. 3. Sunlit-leaf T_c distributions and corresponding thermal mappings.

(a) T_c distribution and fitted GMM

(b) Spatial overlay of detected outliers (red)

Fig. 4. Illustration of GMM-based outlier removal for sunlit-leaf temperatures.

(GMM)¹:

$$p(x) = \sum_{k=1}^2 \pi_k \mathcal{N}(x \mid \mu_k, \sigma_k^2) \quad (2)$$

where π_k are mixing coefficients such that $\sum_{k=1}^2 \pi_k = 1$, and $\mathcal{N}(\cdot)$ denotes a Gaussian distribution with mean μ_k and variance σ_k^2 .

Unlike fixed-threshold methods, which assume unimodal behavior, the proposed GMM allows each image-specific temperature distribution to be represented as a mixture of two components, capturing both typical sunlit-leaf temperatures and extreme thermal responses. The model is fit independently for each image to account for differences in illumination, canopy structure, and segmentation quality.

Outlier detection is performed using a likelihood-based criterion. For each pixel x_i , its likelihood under the fitted mixture is computed as

$$\ell(x_i) = p(x_i).$$

Pixels are classified as outliers if

$$\ell(x_i) < \tau,$$

where τ is defined as the 1st percentile of the likelihood distribution for that image. This criterion identifies low-likelihood tail observations while preserving the main structure of the distribution, including high-temperature values that remain

¹This should not be confused with the earlier GMM-based segmentation benchmark considered in [14].

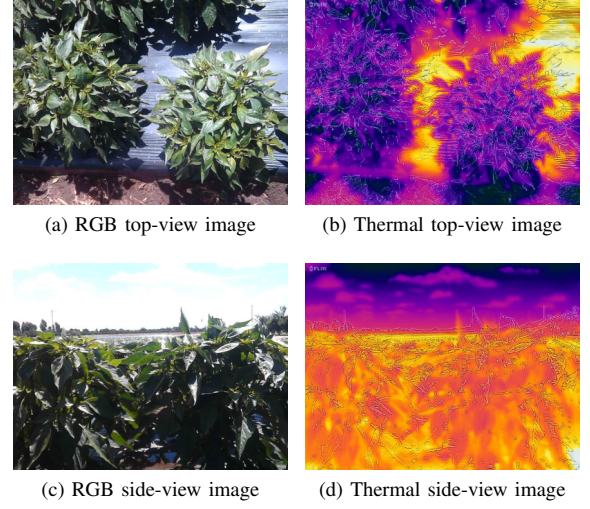


Fig. 5. Example of paired multi-view imaging for the same tomato plant.

consistent with valid sunlit-leaf thermal responses. The results are reported and discussed in Section V-B

H. Impact of Imaging Viewpoint on Sunlit-Leaf Thermal Distribution

To examine the effect of acquisition geometry, we consider a paired multi-view imaging setup. In dense tomato canopies, viewpoint affects the proportion of visible sunlit and shaded leaves, as well as the degree of exposure to background elements such as soil and sky. These factors can systematically alter the thermal distribution of predicted sunlit-leaf pixels and, in turn, bias downstream CWSI estimation.

We therefore capture tomato RGB–thermal image pairs from two viewpoints:

- **Top-view:** the camera is positioned above the canopy.
- **Side-view:** the camera is positioned laterally at canopy height.

All images were acquired on the same day under consistent environmental conditions (clear sky and approximately mid-day). This paired setup enables direct comparison of thermal distributions across viewpoints for the same plant instance. Figure 5 illustrates the experimental setup. Each image is processed using the CIWA pipeline: sunlit-leaf segmentation is first performed on the RGB image, and the resulting mask is then mapped onto the aligned thermal image to extract pixel-wise temperature values corresponding to the predicted sunlit region.

The resulting pixel samples are aggregated by viewpoint (top-view and side-view) to form two empirical distributions representing sunlit-leaf temperature under different acquisition geometries. From these distributions, descriptive statistics including mean, median, standard deviation, skewness, kurtosis, and percentile values are computed. In addition, kernel density estimation (KDE) is used to obtain a continuous approximation of each distribution for visual comparison of their behavior.

This methodology serves as the basis for the viewpoint comparison presented in Section V-C.

V. RESULTS AND DISCUSSION

We present quantitative and qualitative comparisons of CNN- and ViT-based segmentation models on the TMSL21 dataset, followed by an analysis of CWSI-related temperature estimation with GMM-based outlier removal.

A. Model Comparison

Table II summarizes the performance of CNN and Vision Transformer (ViT) models on the TMSL21 tomato sunlit-leaf segmentation dataset. Overall, all architectures achieved similar accuracy, ranging from 0.8100 (DINOv3) to 0.8387 (FC-DenseNet103). However, the models differ more substantially in precision, recall, F1 score, and IoU, which are more relevant for downstream Crop Water Stress Index (CWSI) estimation.

The CNN-based models FRRN-A and FC-DenseNet103 achieved the highest precision scores, 0.7317 and 0.7311 respectively, indicating stronger resistance to false positives and therefore reduced contamination of canopy temperature estimates. Their recall values, however, remain relatively low (0.4522 and 0.4585), suggesting that they follow a conservative prediction strategy that prioritizes reliability over coverage. Compared with prior CIWA results on pistachio crops [14], where FRRN-A and FC-DenseNet103 reached precision values of 0.700 and 0.655 with recall up to 0.510, the tomato setting yields comparable or slightly improved precision but reduced recall, reflecting the increased difficulty posed by dense and highly occluded tomato canopies.

Among the ViT-based models, SegFormer exhibits the most balanced profile, with precision of 0.7018 and the highest recall (0.5024), F1 score (0.5856), and IoU (0.6159). This suggests that transformer-based global context modeling improves the detection of sunlit leaves in dense tomato canopies, where local texture alone is often insufficient. The gain in recall without a substantial loss in precision indicates that SegFormer is better able to resolve ambiguous canopy regions, making it the strongest model overall in terms of segmentation completeness and spatial agreement. Mask2Former shows behavior closer to the CNN-based models, with relatively high precision (0.7244) but more limited recall (0.4619), resulting in only marginal improvements over the CNN baselines. In contrast, DINOv3 and SAM exhibit consistently weaker performance across all metrics, with precision no higher than 0.6504, recall no higher than 0.3593, and IoU of 0.5456. These results suggest that general-purpose or self-supervised foundation models do not transfer effectively to this fine-grained agricultural segmentation task without stronger domain adaptation.

Figure 6 further illustrates model performance through confusion matrices, highlighting the distribution of correctly and incorrectly classified pixels for the sunlit and non-sunlit classes. Figure 7 provides representative visual examples of the predicted masks, allowing qualitative comparison of model behavior relative to the visible-spectrum image, ground-truth label, and overlay.

B. GMM-Based Outlier Removal Results

A consistent issue observed across experiments is instability in sunlit-leaf canopy temperature distributions, primarily caused by extreme thermal values introduced by segmentation noise, boundary leakage, soil contamination, and residual RGB-thermal misregistration. Although these pixels are sparse, they can substantially affect mean T_c because of the heavy-tailed distributions shown earlier in Figure 3a. Applying GMM-based filtering stabilizes the thermal signal by removing statistically inconsistent pixels while preserving coherent sunlit-leaf regions. The proposed trimming procedure yielded highly consistent and low-magnitude removal across the dataset. On average, $1.50\% \pm 0.01\%$ of sunlit pixels per image were identified as outliers. As illustrated in Figure 4, the method removes low-likelihood tail observations while preserving the overall distribution. Importantly, high-temperature pixels are not removed indiscriminately, but are retained when they form coherent components. The spatial overlay further shows that most removed pixels occur near canopy boundaries, where thermal bleeding from soil is more likely. The resulting filtered temperature values are then used for canopy-temperature estimation and downstream CWSI computation. Qualitatively, this step suppresses isolated high- and low-temperature outliers, which are typically concentrated near canopy boundaries, while preserving coherent sunlit-leaf regions that are more likely to represent valid thermal responses.

C. Experimental Results of Imaging Viewpoint Comparison

To quantify the effect of imaging viewpoint, we analyze top-view and side-view sunlit-leaf temperatures as two empirical distributions and compare their descriptive statistics, with particular emphasis on variance and tail behavior, as these are directly relevant to downstream CWSI aggregation. Kernel density estimation (KDE) is used for visualization of the overall distribution structure and does not affect the statistical computations.

Figure 8 shows the corresponding KDE curves. The top-view distribution is wider and exhibits a pronounced right tail, consistent with contamination from non-leaf regions and boundary leakage. In contrast, the side-view distribution remains tightly concentrated around 24–27°C, with minimal tail mass. This heavy-tailed behavior further motivates the need for explicit upper-tail constraints in downstream CWSI calibration and baseline estimation.

The statistical descriptors reported in Table III show that the top-view configuration yields a higher mean canopy temperature (28.02°C) than the side-view configuration (24.98°C), indicating an upward bias caused by preferential sampling of upper-canopy surfaces directly exposed to solar radiation. By contrast, side-view imaging captures a larger proportion of shaded or partially shaded leaves.

This structural difference is reinforced by the dispersion statistics. Top-view distributions exhibit substantially greater variability (standard deviation of 7.65°C) than side-view distributions (2.31°C), indicating a broader and more heterogeneous

TABLE II
TOMATO LEAF SEGMENTATION RESULTS

Architecture	Accuracy	Precision	Recall	F1 Score	IoU
FRRN-A (CNN)	0.8379	0.7317	0.4522	0.5590	0.6036
FC-DenseNet103 (CNN)	0.8387	0.7311	0.4585	0.5636	0.6062
SegFormer (ViT)	0.8385	0.7018	0.5024	0.5856	0.6159
Mask2Former (ViT)	0.8378	0.7244	0.4619	0.5641	0.6058
DINOv3 (ViT)	0.8100	0.6504	0.3546	0.4589	0.5456
SAM (ViT)	0.8144	0.6104	0.3593	0.4523	0.5456

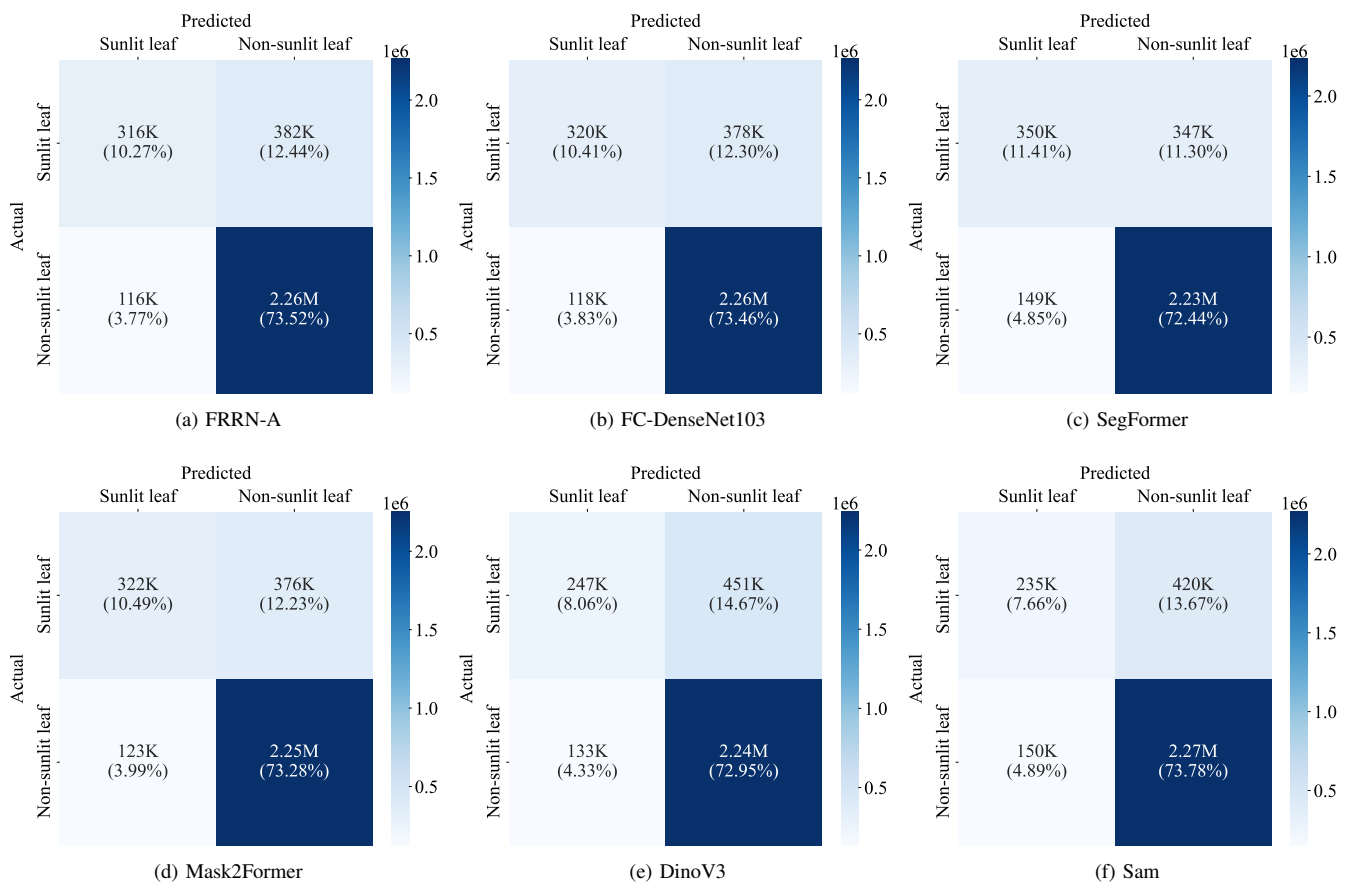


Fig. 6. Confusion matrices (number and percentages of pixels)

TABLE III
TOP VS SIDE IMAGING VIEWPOINT STATISTICS (IN °C)

Viewpoint	Mean	Median	Std	Skewness	Kurtosis	P5	P95	IQR	N Pixels
Top	28.01896	25.90768	7.64908	2.42705	5.42565	21.85497	49.76307	4.06436	700920
Side	24.97923	24.84630	2.30573	0.03914	5.53345	22.52055	28.55015	2.03520	362905

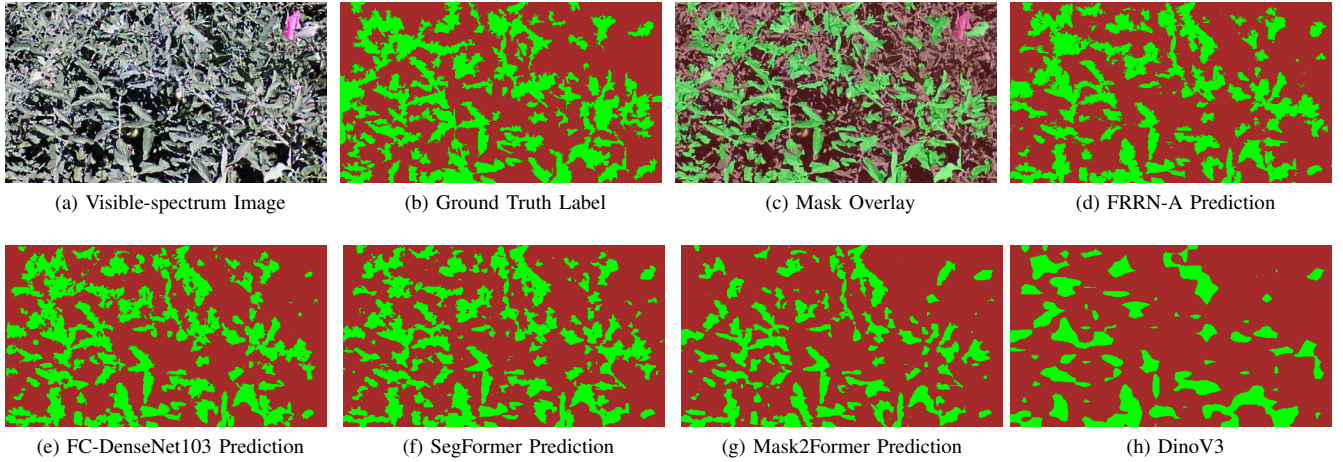


Fig. 7. Representative segmentation results on tomato canopy images across multiple architectures.

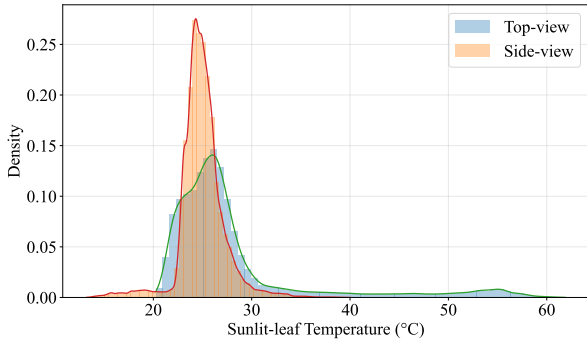


Fig. 8. Kernel density estimation (KDE) of sunlit-leaf temperature distributions for top-view and side-view imagery.

thermal profile. This is consistent with the fact that top-view imagery is more likely to include a mixture of strongly sunlit leaves, partially shaded foliage, and thermally contaminated boundary regions.

The contrast is even more pronounced in the upper tail of the distributions. The 95th percentile of the top-view distribution reaches approximately 49.76°C, whereas the side-view distribution remains bounded near 28.55°C. This suggests that extreme thermal responses are predominantly captured in the top-view configuration. Such behavior is consistent with canopy geometry and with the increased probability of observing both true sunlit hotspots and contaminated pixels arising from soil reflection, segmentation leakage, or residual RGB–thermal misregistration.

Importantly, this difference is not simply a uniform shift of the entire distribution. The top-view distribution has strong positive skewness (2.43), whereas the side-view distribution is nearly symmetric (0.04). This indicates that the divergence is driven primarily by upper-tail amplification rather than by a global temperature offset.

D. Tomato-Specific Baseline for CWSI Calculation

The work of [15] showed that plots of the foliage-air temperature differential ($T_c - T_a$) against vapor pressure deficit

(VPD) yield an approximately linear relationship for plants transpiring at the potential rate during most of the daylight period. This relationship underlies the construction of an environment-normalized CWSI model. In subsequent work, [16] reported corresponding baseline relationships for multiple crop species. One such baseline example is reproduced in Figure 9a.

Following this line of research, we developed a tomato-specific CWSI model for incorporation into CIWA. To do so, we constructed a $T_c - T_a$ versus VPD plot from the thermal data in TMSL21 and applied linear regression to estimate the best-fitting lower baseline. In the CWSI formulation, this regression directly defines the lower baseline term $(T_c - T_a)_{LL}$, which represents the expected temperature difference under non-water-stressed (i.e., fully transpiring) conditions. Specifically, the fitted relationship provides an empirical estimate of the form

$$(T_c - T_a)_{LL} = a + b \cdot \text{VPD}.$$

Before regression, we applied an upper-temperature constraint of 38°C in order to exclude extreme heat values that are unlikely to represent transpiring sunlit-leaf conditions when estimating the lower baseline $(T_c - T_a)_{LL}$. This calibration step is distinct from the GMM-based filtering used to reduce segmentation-related thermal artifacts during canopy temperature estimation. The 38°C cutoff is supported by the distributional analysis in Figure 8, where the KDE curves show a clear separation between the dominant temperature mode (roughly 24–30°C) and a sparse upper tail extending beyond 40°C. For the purpose of baseline estimation, these tail values are treated as inconsistent with the bulk transpiring canopy distribution and are more plausibly attributed to contamination effects than to valid leaf temperatures.

Attempts to fit a baseline directly on the unprocessed data, without excluding the upper tail, produced a positive slope and therefore an implausible physiological relationship between

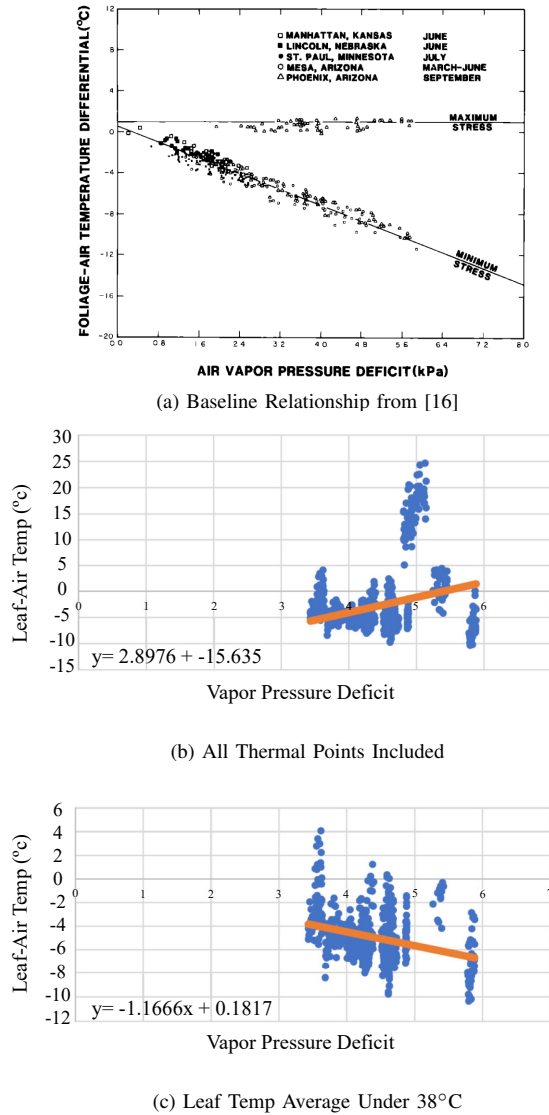


Fig. 9. Leaf-air temperature differential vs. vapor pressure deficit graphs.

$T_c - T_a$ and VPD.² After applying the upper-temperature constraint, the regression yielded a physiologically plausible baseline, as shown in Figure 9c. The tomato-specific CWSI model used in the CIWA web tool was derived from this fitted line, resulting in an estimated slope of -1.1666 and an intercept of 0.1817 .

VI. CONCLUSIONS AND FUTURE WORK

This work extends the Crop Irrigation Web Agent (CIWA) framework to dense tomato canopies for sunlit-leaf segmentation and Crop Water Stress Index (CWSI) estimation. We introduced TMSL21, an annotated RGB-thermal dataset for tomato sunlit-leaf segmentation, and used it to compare CNN-

²For a minimally stressed, fully transpiring plant, the relationship between $T_c - T_a$ and VPD should have a negative slope, as illustrated in Figure 9a. As VPD increases, enhanced transpirational cooling reduces leaf temperature, causing $T_c - T_a$ to decrease.

based and Vision Transformer (ViT)-based models in a dense-canopy setting.

Our results show that reliable CWSI estimation is not determined solely by segmentation accuracy, but by the interaction between segmentation behavior, thermal noise structure, and imaging geometry. Under the standard pipeline, CNN-based models remain the most reliable standalone methods because their higher precision reduces false positives that would otherwise bias canopy temperature. However, when the proposed GMM-based outlier filtering and viewpoint-aware baseline calibration are incorporated, SegFormer emerges as a strong and well-balanced candidate within the full pipeline.

The proposed pipeline further improves robustness by suppressing influential thermal outliers and by accounting for viewpoint-dependent differences in sunlit-leaf temperature distributions. In addition, we derived a tomato-specific CWSI baseline, improving physiological relevance under crop-specific conditions.

Overall, the study advances CIWA toward a more robust end-to-end framework for thermal crop monitoring and CWSI estimation in dense agricultural canopies. Future work will focus on multimodal RGB-thermal models, improved geometric calibration and registration, broader cross-crop validation, and evaluation under real field conditions.

Data Availability

The dataset TMSL21 and the Crop Irrigation Web Agent (CIWA) tool can be found at (www.apanagopoulos.com/ciwa) and https://github.com/AHISLab/ciwa_webapp, respectively.

Acknowledgments

The authors would like to thank Praise Okoli, Quinten dePfyffer, Alexander Madrigal, Nathaniel Madrigal, Dalton Cowin, of the Department of Computer Science at California State University, Fresno; and Arvin Argo, from Fresno Unified School District, for comments and fruitful discussions.

Funding

This research was funded by the F3 Initiative (<https://www.centralvalleycf.org/f3>) and the California State University Agriculture Research Institute (ARI) (<https://csuari.infoready4.com/>).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] Food and U. N. Agriculture Organization. (2012) World agriculture towards 2030/2050: The 2012 revision. [Online]. Available: http://www.fao.org/fileadmin/user_upload/esag/docs/AT2050_revision_summary.pdf
- [2] T. L. Fess, J. B. Kotcon, and V. A. Bedito, "Crop breeding for low input agriculture: a sustainable response to feed a growing world population," *Sustainability*, vol. 3, no. 10, pp. 1742–1772, 2011.
- [3] M. Ozdogan, Y. Yang, G. Allez, and C. Cervantes, "Remote sensing of irrigated agriculture: Opportunities and challenges," *Remote sensing*, vol. 2, no. 9, pp. 2274–2304, 2010.
- [4] L. Konikow. (2013) Groundwater depletion in the united states (1900 to 2008). [Online]. Available: <http://pubs.usgs.gov/sir/2013/5079>

- [5] U. Nations. (2014) Wwdr2014. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000226961>
- [6] J. Rockström, J. Williams, G. Daily, A. Noble, N. Matthews, L. Gordon, H. Wetterstrand, F. DeClerck, M. Shah, P. Steduto, C. de Fraiture, N. Hatibu, O. Unver, J. Bird, L. Sibanda, and J. Smith, "Sustainable intensification of agriculture for human prosperity and global sustainability," *Ambio*, vol. 46(1), pp. 4–17, 2 2017.
- [7] E. A. Abioye, M. S. Z. Abidin, M. S. A. Mahmud, S. Buyamin, M. H. I. Ishak, M. K. I. Abd Rahman, A. O. Otuoz, P. Onotu, and M. S. A. Ramli, "A review on monitoring and advanced control strategies for precision irrigation," *Computers and Electronics in Agriculture*, vol. 173, p. 105441, 2020.
- [8] R. D. Jackson, S. B. Idso, R. J. Reginato, and P. J. Pinter, "Canopy temperature as a crop water stress indicator," *Water Resources Research*, vol. 17, no. 4, pp. 1133–1138, 8 1981.
- [9] V. Alchanatis, A. Cohen, Y. Cohen, O. Levi, and A. Naor, "Multi-modal remote sensing for enhancing detection of spatial variability in agricultural fields," in *Spatial2 Conference: Spatial Data Methods for Environmental and Ecological Processes, Foggia (IT), 1-2 September 2011*. Università degli studi di Bergamo, 2011, pp. 1–4.
- [10] B. Gardner, D. Nielsen, and C. Shock, "Infrared thermometry and the crop water stress index. ii. sampling procedures and interpretation," *Journal of production agriculture*, vol. 5, no. 4, pp. 466–475, 1992.
- [11] D. Cui, Q. Zhang, M. Li, G. L. Hartman, and Y. Zhao, "Image processing methods for quantitatively detecting soybean rust from multispectral images," *Biosystems engineering*, vol. 107, no. 3, pp. 186–193, 2010.
- [12] Y. Dai, R. E. Dickinson, and Y.-P. Wang, "A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance," *Journal of climate*, vol. 17, no. 12, pp. 2281–2299, 2004.
- [13] Z. Wang, L. Liu, B. Wang, and X. Wang, "Aerosol increases both sunlit and shaded leaf photosynthesis rate but with different mechanisms," in *AGU Fall Meeting Abstracts*, vol. 2017, 2017, pp. B51F–1879.
- [14] M. Pantelidakis, A. A. Panagopoulos, K. Mykoniatis, S. Ashkan, R. C. Eravi, V. Pamula, E. C. Verduzco III, O. Babich, O. P. Panagopoulos, and G. Chalkiadakis, "Identifying sunlit leaves using convolutional neural networks: An expert system for measuring the crop water stress index of pistachio trees," *Expert Systems with Applications*, vol. 209, p. 118326, 2022.
- [15] S. B. Idso, R. D. Jackson, P. J. Pinter, R. J. Reginato, and J. L. Hatfield, "Normalizing the stress-degree-day parameter for environmental variability," *Agricultural Meteorology*, vol. 24, pp. 45–55, 1981.
- [16] S. B. Idso, "Non-water-stressed baselines: a key to measuring and interpreting plant water stress," *Agricultural Meteorology*, vol. 27, p. 59–70, 1982.
- [17] X. Wang, W. Yang, A. Wheaton, N. Cooley, and B. Moran, "Automated canopy temperature estimation via infrared thermography: A first step towards automated plant water stress monitoring," *Computers and Electronics in Agriculture*, vol. 73, pp. 74–83, 07 2010.
- [18] M. H. Calp and V. K. Solanki, "Classification of plant species with iris dataset using ann, knn and k-means algorithms," in *Proceedings of the Eighth International Conference on Research in Intelligent Computing in Engineering*, ser. Annals of Computer Science and Information Systems, P. Kumar, M. Cardona, V. K. Solanki, T. D. Tan, and A. Wahid, Eds., vol. 38. PTI, 2023, p. 7–10. [Online]. Available: <http://dx.doi.org/10.15439/2023R63>
- [19] M. Arsov, E. Zdravetski, P. Lameski, R. Corizzo, N. Koteli, K. Mitreski, and V. Trajkovic, "Short-term air pollution forecasting based on environmental factors and deep learning models," in *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems*, ser. Annals of Computer Science and Information Systems, S. Agarwal, D. N. Barrell, and V. K. Solanki, Eds., vol. 21. IEEE, 2020, pp. 15–22. [Online]. Available: <http://dx.doi.org/10.15439/2020KM211>
- [20] Flir. (2022) Flir ax8 datasheet. [Online]. Available: <https://support.flir.com/DsDownload/Assets/71201-0101-en-US.html>
- [21] A. dos Santos Ferreira, D. M. Freitas, G. G. da Silva, H. Pistori, and M. T. Folhes, "Weed detection in soybean crops using convnets, computers and electronics in agriculture," vol. 143, pp. 314–324, 2017.
- [22] T. Pohlen, A. Hermans, M. Mathias, and B. Leibe, "Full-resolution residual networks for semantic segmentation in street scenes," in *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, 2017.
- [23] S. Jégou, M. Drozdal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 11–19.
- [24] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "Segformer: Simple and efficient design for semantic segmentation with transformers," *arXiv preprint arXiv:2105.15203*, 2021.
- [25] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar, "Masked-attention mask transformer for universal image segmentation," 2022. [Online]. Available: <https://arxiv.org/abs/2112.01527>
- [26] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3213–3223.
- [27] B. Zhou, H. Zhao, X. Puig, T. Xiao, S. Fidler, A. Barriuso, and A. Torralba, "Semantic understanding of scenes through the ade20k dataset," 2018. [Online]. Available: <https://arxiv.org/abs/1608.05442>
- [28] N. Carion, L. Gustafson, Y.-T. Hu, S. Debnath, R. Hu, D. Suris, C. Ryali, K. V. Alwala, H. Khedr, A. Huang, J. Lei, T. Ma, B. Guo, A. Kalla, M. Marks, J. Greer, M. Wang, P. Sun, R. Rädle, T. Afouras, E. Mavroudi, K. Xu, T.-H. Wu, Y. Zhou, L. Momeni, R. Hazra, S. Ding, S. Vaze, F. Porcher, F. Li, S. Li, A. Kamath, H. K. Cheng, P. Dollár, N. Ravi, K. Saenko, P. Zhang, and C. Feichtenhofer, "Sam 3: Segment anything with concepts," 2025. [Online]. Available: <https://arxiv.org/abs/2511.16719>
- [29] O. Siméoni, H. V. Vo, M. Seitzer, F. Baldassarre, M. Oquab, C. Jose, V. Khalidov, M. Szafraniec, S. Yi, M. Ramamonjisoa, F. Massa, D. Haziza, L. Wehrstedt, J. Wang, T. Darcet, T. Moutakanni, L. Sentana, C. Roberts, A. Vedaldi, J. Tolan, J. Brandt, C. Couprie, J. Mairal, H. Jégou, P. Labatut, and P. Bojanowski, "Dinov3," 2025. [Online]. Available: <https://arxiv.org/abs/2508.10104>
- [30] L. Prechelt, "Early stopping-but when?" in *Neural Networks: Tricks of the trade*. Springer, 1998, pp. 55–69.
- [31] A. Leigh, J. D. Close, M. C. Ball, K. Siebke, and A. B. Nicotra, "Leaf cooling curves: measuring leaf temperature in sunlight," *Functional Plant Biology*, vol. 33, no. 5, pp. 515–519, 05 2006. [Online]. Available: <https://doi.org/10.1071/FP05300>
- [32] J. D. Muller, E. Rotenberg, F. Tatarinov, I. Vishnevetsky, T. Dingjan, A. Kribus, and D. Yakir, "'dual-reference' method for high-precision infrared measurement of leaf surface temperature under field conditions," *New Phytologist*, vol. 232, no. 6, pp. 2535–2546, 2021. [Online]. Available: <https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.17720>