Non-destructive Leaf Area Index estimation via guided optical imaging for large scale greenhouse environments

メタデータ	言語: en
	出版者: Elsevier
	公開日: 2024-04-04
	キーワード (Ja):
	キーワード (En): Tomato, Deep learning, LAI, greenhouse farming, Agriculture
	作成者: Baar, Stefan, 小林, 洋介, Horie, Tatsuro, 佐藤, 和彦, 須藤, 秀紹, 渡邉, 真也
	メールアドレス:
	所属: 室蘭工業大学, 室蘭工業大学
URL	http://hdl.handle.net/10258/0002000066

Non-destructive Leaf Area Index estimation via guided optical imaging for large scale greenhouse environments

Stefan Baar[1], Yosuke Kobayashi[1], Tatsuro Horie[2], Kazuhiko Sato[1], Hidetsugu Suto[1], Shinya Watanabe[1]

[1] *Muroran Institute of Technology, Muroran, Hokkaido, Japan*

[2] *Air Water Co.Ltd. Chitose, Hokkaido, Japan.*

Abstract

This paper presents a financially viable and non-destructive rail-based video monitoring method that utilizes optical image segmentation to estimate the canopy leaf area index (LAI) of greenhouse tomato plants. The LAI is directly related to the time-dependent crop growth and indicates plant health and potential crop yields. A rail-guided mobile camera system was commissioned that records continuous images by scanning multiple rows of two tomato plant species for over two years. UNET semantic image segmentation of the individual image frames was performed to compute the relative leaf area over time. This study also describes the image annotation process necessary to train the neural network and evaluate the segmentation results. The results are calibrated and compared to the defoliation-based (destructive) LAI estimation performed by the grower. This UNET segmentation performs well, which is enabled through the controlled environment and the well-defined boundary conditions provided by the greenhouse environment and the managed measurement conditions. Our results deviate from the manual LAI estimation by less than ten percent. Further, we are able to minimize confusion between foreground and background plants and other obstructions with an estimated error smaller than three percent, which is strictly necessary to produce reproducible results.

Keywords: tomato, deep learning, LAI, greenhouse farming, agriculture

1. Introduction

Recently, the management of large-scale greenhouse farming environments, as well as phenotyping, has attracted the interest of many machine learning and deep learning researchers. In general, modern machine learning methods are necessary to meet agricultural production challenges related to sustainability, food security, environmental conservation, and productivity [1, 2]. From a physical perspective, even a well-managed greenhouse ecosystem is complex and difficult to predict with many non-linear interrelationships arising from the chaotic nature of plants and the outside environment [3]. Phenotyping is critical in greenhouse environmental management for pest and disease detection and generally to monitor plants for predicting the development of dry yield.

In particular, when monitoring the general health of vines and optimizing the monthly fruit harvest yield, the Leaf Area Index (LAI), is the main indicator for managing greenhouse environments and controlling defoliation [4, 5]. The plant canopy density and the general number of leaves affect the energy, hydration [6, 7], and overall CO_2 balance within a greenhouse through transpiration [8]. Crop leaf growth strongly affects the assimilation capability of photosynthetically active radiation [9]. Furthermore, [10] showed that the fraction of light interception (I) is connected to the LAI through the following power law:

$$I = 1 - e^{-kLAI}. \tag{1}$$

The LAI is a dimensionless quantity defined as the one-sided leaf area (A_L) per reference area, where the reference area (A_R)can be considered as a square hull (first-order convex hull) around the leaf.

$$LAI = A_L/A_R. \tag{2}$$

It should be noted that this definition is only valid for broad leaf canopies, as presented by [11] and can be more complex when including multiple leaf layers within the plant canopy[12]. This means that the LAI could be larger than one for staged and overlapping leaves, depending on the direction of the incident light. Previous research has shown that the LAI is a comprehensive indicator of variations in environmental, biological, and plant-structural conditions. Because the time-dependent LAI reflects the CO_2 content and material cycles in the plant canopy [3], one can use the time-dependent LAI to control plant characteristics, for example, plant height and fruit/plant growth, through defoliation [13].

Variations in the observed LAI value for one plant at a precise time, critically depend on the measurement method, environmental conditions, and theoretical completeness of the measurement approach. In principle, there are two groups of measurement methods: active/destructive and passive/non-destructive LAI measurements, many of which have recently been summarized by Popovic et. al. [14]. The most common and simplest method is active plant measurement by defoliation, which requires the grower to regularly remove all the leaves in a predetermined reference space and count the number of leafs and area (e.g., measuring the breadth and length of individual leaves) manually or by computer analysis, e.g., color [15, 16, 17] or otsu [18, 19] segmentation, etc. There are numerous non-destructive remote sensing approaches that depend on specific environmental conditions. They are either radio-based approaches, as elaborated by Orlando et al. [20] and Campos et al. [21], image-based approaches [11] or 3D sensing approaches such as photogrammetry [22] point-cloud based evaluation [23]. For most image-evaluating solutions, rule or deep learning-based segmentation techniques are used to separate the relative leaf area in the image from the background and other distractors. This is particularly a trend in hand-held applications, as found in [24, 25, 16], and [20], where either the relative leaf area (in pixels) or the leaf-to-leaf voids are calibrated to reflect the LAI amplitude. Computing LAI is limited by the precision and accuracy of the detection/segmentation method. However, it is seldom noted that reproducibility and precision are highly dependent on the quality of the segmentation routine, especially among the approaches that generally discount spatial and contextual considerations. In other words, camera position changes, field of view (FOV, parts of the plant that are imaged), confusion with other plants, and distractors can greatly influence the observational accuracy.

Recently Fang et. al. [26, 27] have confirmed relatively high variations in reproducibility when estimating *LAI* in image based smartphone apps, which are most likely caused by environmental and plant-morphological changes over time and the limited environmental perception abilities of the individual app based segmentation routines. To improve the *LAI* estimation accuracy and reproducibility, it is necessary to establish a robust *LAI* measurement routine. Advanced deep learning based segmentation methods have great potential to improve segmentation accuracy [11, 28] and introduce environmental perception, which has been widely used in autonomous vehicle and drone research [29, 30, 31, 32].

In this study, we present a novel approach to estimate the *LAI* of greenhouse grown vines of Solanum lycopersicum also known as the common Tomato plant. Our approach is non destructive, reproducible and produces results comparable to manual *LAI* estimation through defoliation, while being more efficient. The measurement method is optimized for large greenhouse facilities with long and narrow plant rows, where a camera (optical RGB) mounted on a rail wagon is scanning an entire row with *LAI* being computed from the neural network based segmentation map of the produced image data.

We estimate *LAI* by analyzing video frames captured using a rail-mounted camera system. In the next section, Methods and Materials, we introduce the Target greenhouse environment, and described how the reference data is obtained. The optical LAI measurement approach is introduced and explained in the section 2.2. The Results section compares this method with traditional *LAI* estimation through defoliation. Finally, general findings and the validity of the approach are discussed in section, Discussion and Conclusions.

2. Methods and Materials

2.1. Target greenhouse environment

This section introduces the target greenhouse environment and the technique used to determine LAI_{grower} as a reference for the non-destructive LAI measurement method. This research was conducted at the Air Water Greenhouse Complex in Chitose, Hokkaido, Japan.

Hokkaido is the northern-most of Japan's islands and lies in a temperate climate zone. This is in strong contrast to the rest of Japan, which is classified as subtropical. More specifically, the area of central Hokkaido is classified as plant hardiness zone 5, which means temperatures in winter can reach values below $-20\,°C$ and are usually not above $30\,°C$ in summer . Therefore, vines, such as tomatoes, planted in greenhouse environments are usually grown during summer, starting in February and ending in November.

The target greenhouse covers an area of 4 ha (40,000 m^2) containing approximately 50,000 tomato plants. In this study, the LAI is evaluated for the mini-tomato bearing daltary and mid-size tomato bearing tomimaru species. Until the development of the proposed deep learning-based procedure, the grower estimated the LAI by partially defoliating the canopy of a predefined set of reference plants. In this paper, this method of LAI measurement is termed grower LAI (LAI_{grower}). The LAI_{grower} estimation for both tomato species was performed by measuring the leaf attributes of four reference tomato plants in the center of the tomato greenhouse and computing the LAI as follows:

$$LAI_{grower} = \alpha_{D|T} \times S_L \times N_L \tag{3}$$

where S_L is the area of the square enclosing a single leaf, computed as the product of the leaf width w_L the leaf height h_L in m^2. N_L: Nr. of Leaves per m^2, α: The species-dependent absorption coefficient, which depends on the leaf shape. This value is provided by the distributer of the tomato seeds and

[1]Chitose climate according to weather park (`https://ja.weatherspark.com`)

presents the mean fraction of the leaf area to its enclosing square computed by the leaf length and width. It roughly accounts for the complex shape of each leaf, from which the average leaf area (A_L) can be computed as follows:

$$A_L = \alpha_{D|T} \times \sum_{n}^{N_L} \frac{w_{Ln} \times h_{Ln}}{N_L} = \alpha_{D|T} \times S_L. \quad (4)$$

The values for α for tomimaru and daltary are $\alpha_T = 0.620$ and $\alpha_D = 0.618$, respectively. The different values of α for tomimaru and daltary are the result of the slightly varying leaf morphology of the two tomato species. It should be noted that only the leaves of the upper branches of every plant were considered in this study.

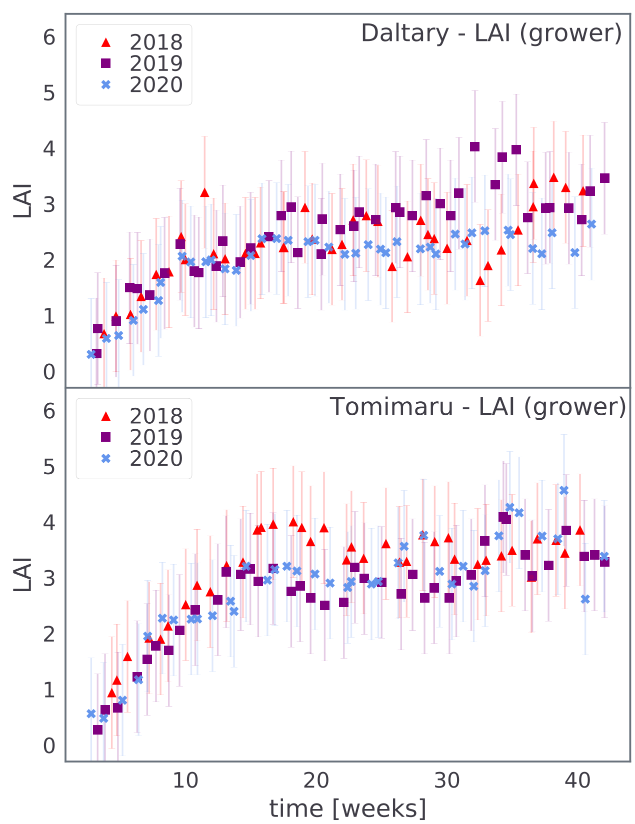

Figure 1: Grower LAI for both tomato species, daltary (top) and tomimaru (bottom) for the years 2017 to 2020.

The LAI_{grower} evolution from 2017 to 2020 is presented in Figure 1. It should be noted that the uncertainties of each LAI measurement were relatively high because of the reference plants were only partially defoliated and the fact that there is only a small ensemble of reference plants used in this study. The seedlings were planted out in cultivation lines, which were spaced approximately 1.2 meters apart. The canopy tip was held by a guide string to secure and displace the plant canopy during the maturity process. Displacement in combination with defoliation below the fruit line is necessary to keep the canopy and fruit at a constant height, which is convenient for harvesting, plant management and the UNET-based LAI estimation approach, which is discussed in the next section.

2.2. Optical LAI estimation (through UNET)

[2] Throughout this paper, the week date is referenced as the ISO week number.

2.2.1. Methodology

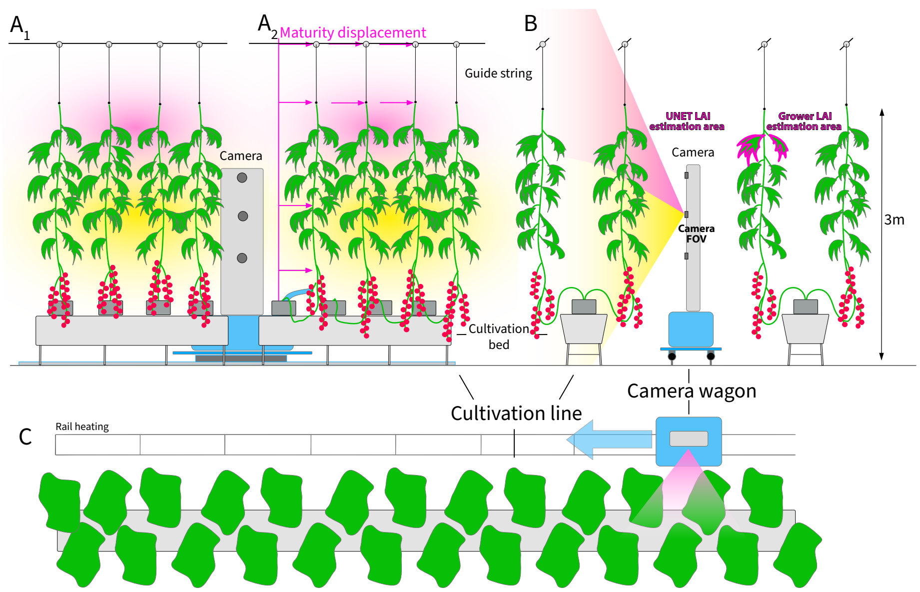

Figure 2: Data collection environment and cultivation line. Illustration of the cultivation line with young tomato plants (A_1), mature tomato plants right (A_2), and camera car on the back. Canopies are kept vertically in place because of the regular defoliation and displacement of the guide string. B: side view of the cultivation line with the camera car and the camera's FOV indicator. The areas used to estimate LAI_{grower} and the proposed UNET approach are indicated in pink. C: Top view of the cultivation line with camera car's direction of travel. The rail is used to guide the camera car as well as for heating in spring and fall.

This section describes our non-destructive LAI estimation approach as well as the data processing, calibration, and evaluation procedures. The proposed LAI estimation approach relies on a moving camera system, which scans an entire tomato plant row and creates a relatively large number of frames for various tomato plants from various angles using a large FOV. A camera with a constant elevation is mounted on a rail wagon and used to scan one row of tomato plants, as presented in Figure 2.

The camera is at a constant distance from the tomato row and has a constantly wide spherical FOV to keep more than 90% of the canopy within the FOV. Furthermore, all the tomato rows in the greenhouse had approximately the same width and height, with all of them filling a similar cylindrical volume. The elevation angle of the lens is zero. Automatic exposure compensation used with an exposure time range between (0.0002 - 0.001)s. The running velocity is predetermined by the manufacturer. The camera captured video frames at an adequate rate of 10 fps, while the wagon traverses the rail at a constant velocity of approximately $v = 0.2m/s$. Further, we have confirmed the velocity through independent distance and time measurements. Owing to the wide lens angle, plant images are captured from multiple perspectives. In principle, this improves the differentiation between the tomato plant rows in the front and those in the background because the rows traverse the FOV with different relative velocities relative to the wagon.

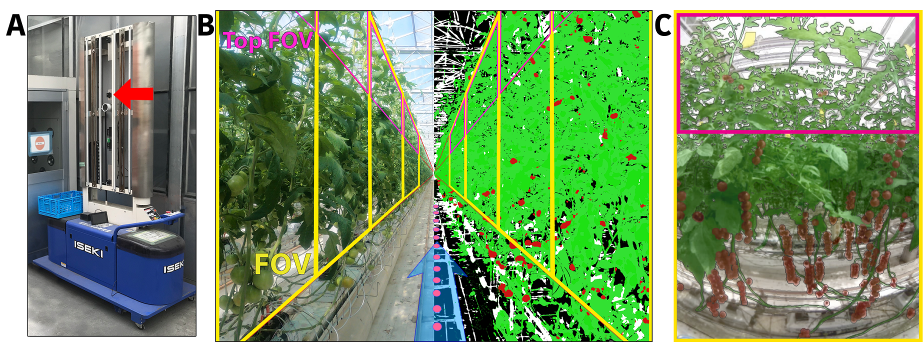

Figure 3: A: Camera mounted wagon: The camera is indicated by the red arrow. B: Image of greenhouse tomato row with actual image (left) and annotation overlay (right) with background (black), building structure (white), plant (green), and tomato (red). C: The tomato row image as observed by using the annotation overlay. The yellow frame indicates the FOV of the image. The pink frame indicates the area in which LAI_{grower} is measured.

The image capturing procedure is demonstrated in Figure 3, where image A shows the camera-mounted wagon, and image B is the tomato plant row (left) with the UNET segmentation overlay (right) and the frame FOV (yellow and pink outlines). Image C shows the plant row, as seen by the camera system. In this project, the full FOV (yellow) and the upper one-third of the field of view (pink) were evaluated. This is necessary to calibrate and compare the result based on the upper canopy FOV to LAI_{grower} and because the LAI_{grower} data could be obtained for only the upper canopy containing the top branch of the tomato plant, as presented in Figure 2 B.

2.2.2. Data processing overview

A general map of how the time-dependent LAI was estimated from the video data is presented in Figure 4. The proposed method relies on two procedures: 1) image segmentation and 2) calibration to compute the LAI from the projected leaf area. The assumption that the LAI can be sufficiently estimated from the projected leaf area observed by the camera is critical for justifying the proposed approach and is elaborated in the following section.

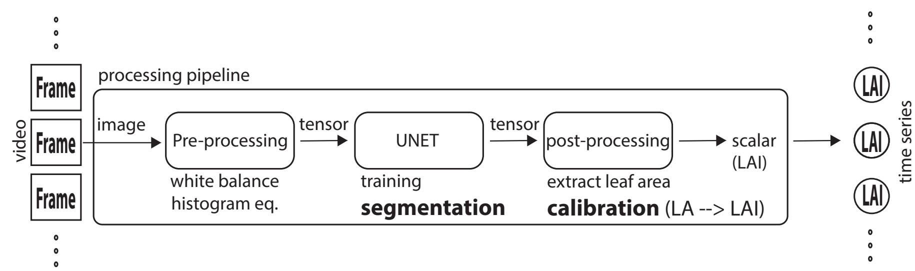

Figure 4: Data processing pipeline. Illustration of how the LAI time-series data are extracted from the image data.

2.2.3. LAI calibration

Owing to the non-destructive nature of the proposed sensing approach and the complex three-dimensional morphology of the canopy, the absolute value for the LAI cannot be computed directly. Therefore, the time-dependent variability in the LAI can only be obtained by assuming that the obtained relative leaf area (A_L) is proportional to the LAI in time and space.

$$\frac{dLAI(t)}{dt} = \frac{dA_L(\vec{r},t)}{dt} \tag{5}$$

where $\vec{r}$ represents the spatial vector from which the canopy is observed, and A_L is the fraction defined by the ratio of pixels containing leaves to the total number of pixels in the image. Because the FOV of the camera is sufficiently large to ensure that the entire canopy is within the FOV at any time t, a linear relationship between A_L and the LAI can be assumed, introducing a gain calibration (g) as follows:

$$LAI(t) = g \times A_L(t) \tag{6}$$

The expected gain amplitude can be determined simply by computing g for a randomly selected set of times T as follows:

$$\bar{g} = \sum_{t}^{T} \frac{LAI(t)}{A_L(t)} / T \tag{7}$$

Furthermore, the standard deviation of g can provide a measure of the extent to which the trend of A_L deviates from the directly measured trend of the LAI. To obtain a good estimation of A_L, each image was segmented for deep learning and the number of pixels associated with A_L was computed as presented in the next section.

2.2.4. Deep learning based image segmentation

There are several semantic segmentation models of which the most prominent are the encoder-decoder convolutional neural networks such as Inverse Hour Glass and UNET. These have been widely discussed over the last decade and are well utilized in the fields of medical and microbiological research [33, 34, 35, 33]. Trained Encoder-decoder network models fixed to the image resolution and its multiples. However the results generated by those networks are generally do not dependent on the resolution of the input image, as long as it fits into the model and as long as it is resized to match the size of the training images. However, the individual feature scale within the image needs to be preserved. This is favorable for our approach, since the feature size within the image das not vary strongly. A costume UNET model was constructed using PyTorch [36] and used

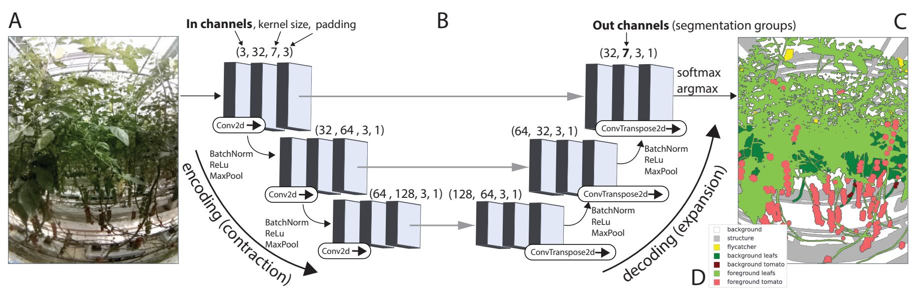

Figure 5: Image segmentation procedure. A: Three-channel input image. B: UNET neural network with three contraction/extraction layers for training and inference. C: Output mask containing seven segmentation groups labeled in D.

to performed semantic segmentation to resolve the seven segmentation groups, as presented in Figure 5 D. The UNET model consists of three contraction and expansion blocks with dimensions ranging from 32 to 128[1].

The input consists of a $3 \times 640 \times 640$ (RGBxXxY) image tensor, which is normalized to contain float32 values between zero and one. The trained model returns a tensor with dimensions of $(640 \times 640 \times 7)$ containing seven intensity maps (one for each group). Using the softmax function, the intensity maps are then converted into a probability map from which the resulting annotation mask is computed using the argmax function. Of the seven index groups, only the index groups describing the background and foreground leaf area were strictly necessary. Confusion between leaves and other objects is increased when only a limited number of annotation samples are available and can be avoided by increasing the number of annotation groups at the cost of requiring increased training (in the form of time and computational resources).

[1] Further information regarding the UNET model and the data processing approach are available through the github repository https://github.com/StefanBaar/LAI_network.

3. Data analysis and Results

This section describes the automated data preparation, analysis, and post-processing pipeline presented in Figure 4, as well as the annotation procedure and the results. The video data consists of compressed mp4 files with an average length of approximately two thousand frames recorded at ten fps. Each video frame has been white balance corrected based on the color information of the highlights (pixels within a five-pixel radius around saturated pixels) of the upper third of each frame. Furthermore, histogram equalization was performed using 99% of the RGB histogram. The preprocessed images were then segmented into seven annotation groups and the number of pixels associated with the leaf area for each frame was saved. Next, the annotation process is described; it produces the annotation groups and their pixel-based locations on which the model is trained.

3.1. Annotation and augmentation

Pixel-based annotations were prepared using Adobe Photoshop because it is easy to use and supports drawing tablets. One layer was used for each group in addition to the original image used as the first layer in the layer stack. Annotation examples are presented in Figure 6, where the raw images are shown side-by-side with their respective annotations Routines were prepared to convert psd files into Pytorch tensors using the Python library psd-tools [37]. Two hundred annotations were prepared that contained randomly selected pseudo images from the individual tomato datasets. The images and annotations were augmented using a random crop (70%-100% image area), random rotation, and random horizontal spatial flipping. The training data set was inflated to 10k images through spatial augmentation using the scipy [38], scikit-image [39], and PyTorch [36] sub-modules, as well as costume functions. Furthermore, on-the-fly random augmentations were performed during training using random brightness adjustment (± 20%), random histogram equalization (0% - 2%), and a random color jitter (amplitude: ± 10%).

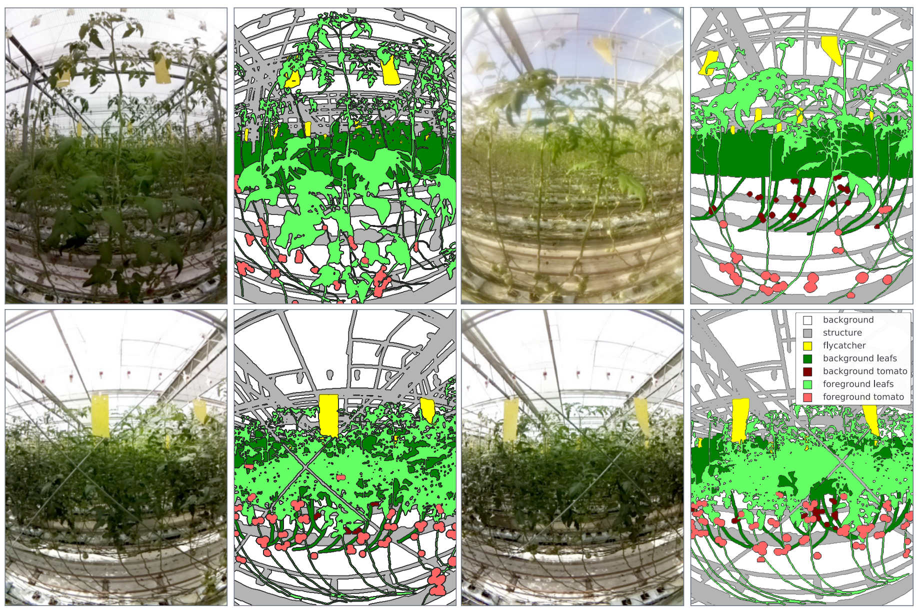

Figure 6: Annotation overview. Four annotation examples with each pre-processed (histogram equalized and white balanced) raw image on the left and the representative annotations on the right.

3.2. *Training*

The augmented dataset was divided into equally sized training and evaluation datasets. The network was trained using the ADAptive Moment (ADAM) estimation optimizer to compute the gradient descent and update the weights of the UNET model during training. ADAM was chosen because it is generally known for its computational efficiency, minimal memory requirements, appropriateness for noisy and sparse gradients, and that it is well implemented in PyTorch [36]. As a loss function, PyTorch implementation was used for cross-entropy loss. A variable learning rate was used depending on the training progress, $lr = (0.003, 0.002, 0.001, 0.0008)$ and the network was trained for 550 epochs with a batch size of 40. Training was implemented in parallel over four NVIDIA GTX 1080Ti GPUs until the training and validation errors reached values below 10% and 20%, respectively, for at least 10 epochs, as presented in Figure 7.

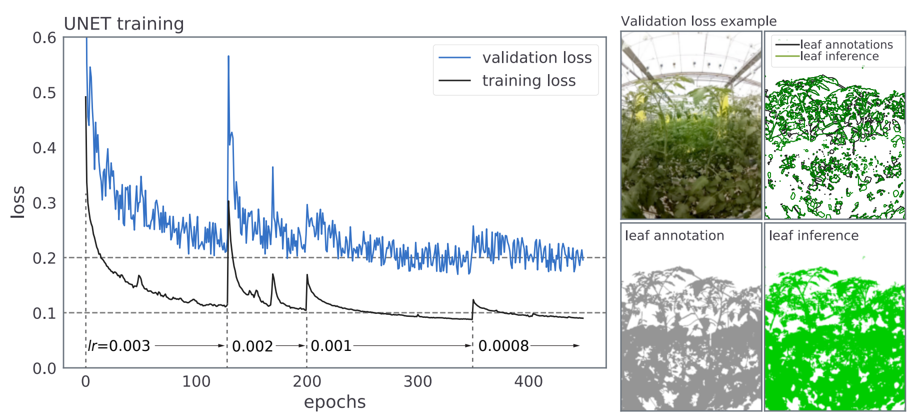

Figure 7: Left: training and validation loss for each epoch. Right: Image example comparing leaf annotation and leaf inference.

3.3. *Results*

This section presents the evaluation of the LAI estimation of mini-tomato bearing daltary and mid-size tomato bearing tomimaru plants from May 2019 to July 2021 and the results are compared to a defoliation-based LAI measuring approach conducted by the grower. Both the relative leaf area A_L and LAI_{grower} were measured for the upper one-third of the FOV covering two types of tomato plants, as presented in Figure 3 B. A number of segmentation results produced by inference is shown in Figure 8.

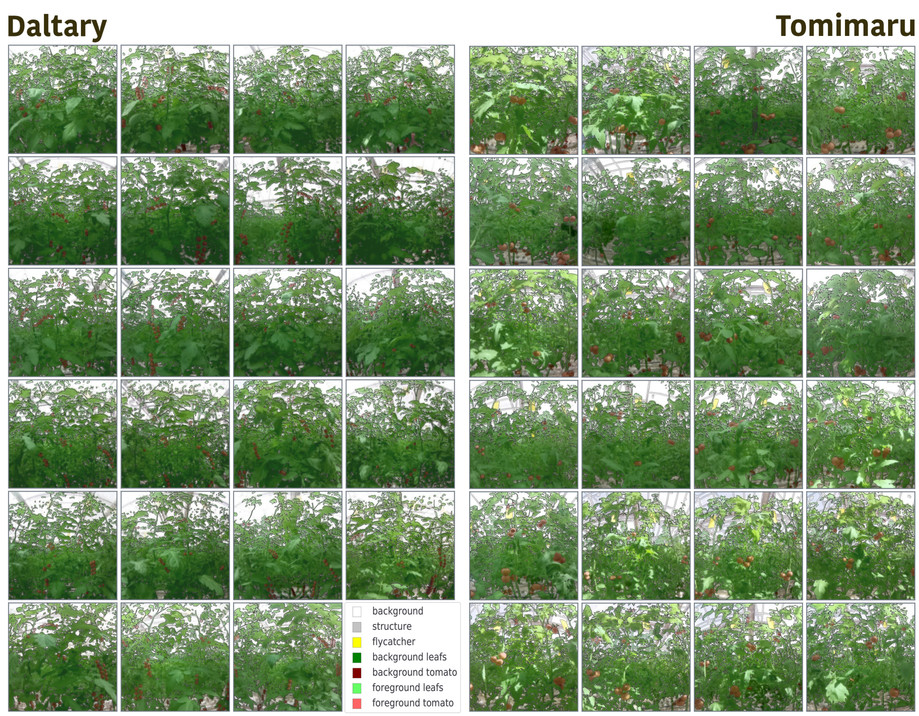

Figure 8: Segmentation overview: Forty-seven segmentation examples created via the inference of raw images with the proposed UNET model for the tomato species daltary (left) and tomimaru (right). Each example contains a pre-processed (histogram equalized and white balanced) raw image, as well as a superimposed annotation mask.

We present the correlation between inference and annotation area of foreground and background leafs (as fraction of image area) in Figure refforeback for the validation dataset of the individual tomato species and their combination.

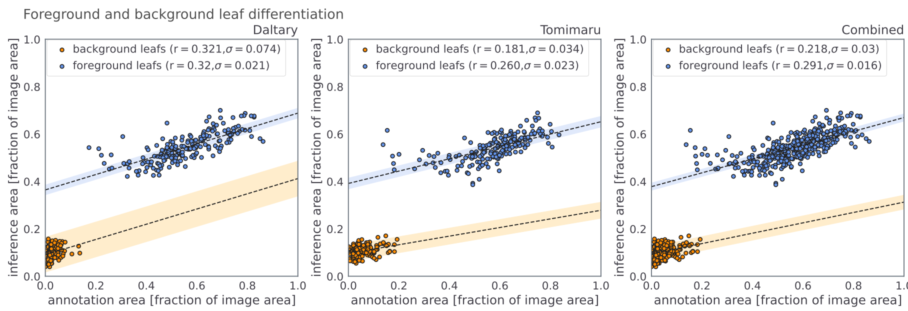

Figure 9: Confidence of inferred background and foreground area in relation to their respective annotation areas for two hundred images of daltary (left), tomimaru (center), and both species combined (right).

The figure shows, that for both species, foreground and background are well differentiated. Also the standard deviation σ is small, which hints towards a strong proportionality between between inference and annotation area. As expected, the correlation coefficient (r) is relatively low, due to the fact that annotation area is determined by the perimeter of each annotation group. One might note from the right hand side of Figure 7, that while the general morphology of the inference map is very similar to the annotation map, the residual between the two maps is not zero. This is manly due to annotation offset and is not problematic for our approach. This is because, we do not consider the absolute area, but the relative area, which is highly correlated between annotation and inference maps, indicated by a small $\sigma < 0.03$ for the foreground leaf map. While the foreground leaf area is used to estimate LAI, the background leaf area is omitted. Its purpose lies solely in eliminating false positives.

As previously mentioned, only the relative leaf area that is proportional to the LAI can be measured by gain amplitude $g(t)$, which is computed using the relative leaf area and LAI_{grower} of the year 2019. The time variation of $g(t)$ for the two tomato species is shown in Figure 10.

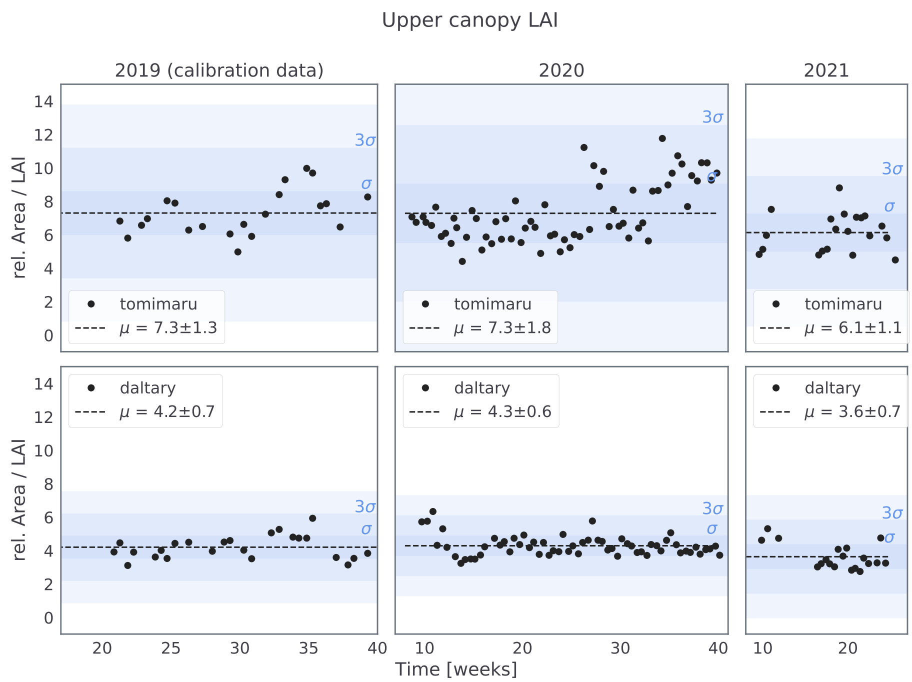

Figure 10: Gain calibration amplitude (relative leaf area/grower LAI) for the tomimaru (top) and daltary (bottom). The dashed line denotes the mean calibration amplitude. The bars denote the variability of the measurements.

The mean of the time-dependent g was computed to calibrate the UNET LAI for the years 2020 and 2021 for the upper third of the FOV, as presented in Figure 11. The deviation between the UNET-based and defoliation-based LAI estimation method was computed to be less than ten percent. The different gain amplitudes for tomimaru and daltary arise from their different leaf shapes.

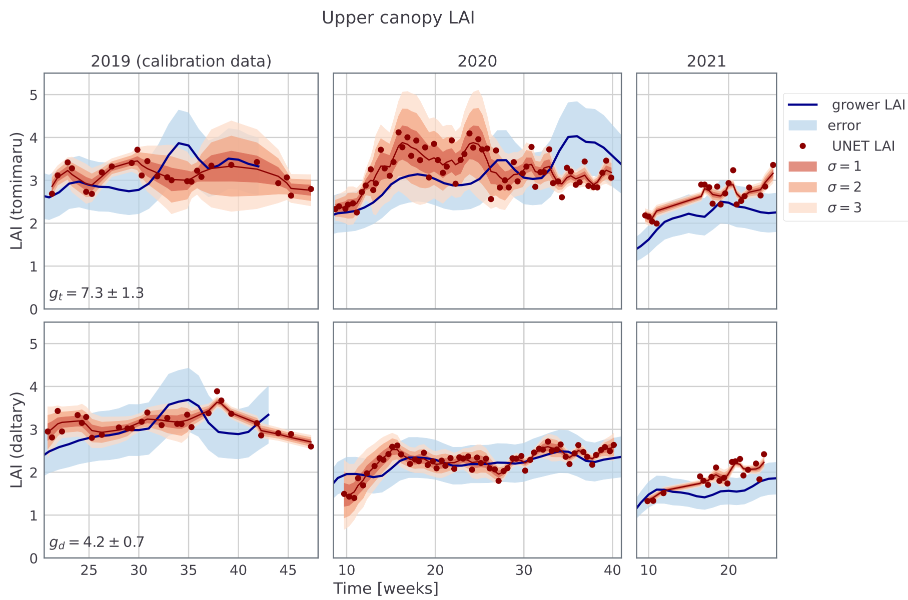

Figure 11: The **LAI of the upper plant canopy** determined by the grower (blue line) and computed through the UNET semantic segmentation (red dots) with its time averaged spline interpolation (red line) for the times span between May 2019 and June 2021. The variability σ of the measurement data (2000 segmentation frames per observation) in shades of red.

The $LAIs$ for the tomimaru and daltary tomato plants were estimated from the gain correction amplitude g (computed from the upper FOV and the relative leaf area of the entire FOV). Figure 12 presents the average per plant LAI time evolution. The full canopy LAI data cannot be directly compared because defoliation-based LAI measurements could not be made for the entire plant canopy since it would have been time-consuming and labor-intensive. However, the general trend is as expected by the grower and exhibits a somewhat flatter tendency when compared to the upper canopy LAI. This is because the branches below the fruit line are pruned regularly, which causes only a slight variance in canopy leaf numbers over time.

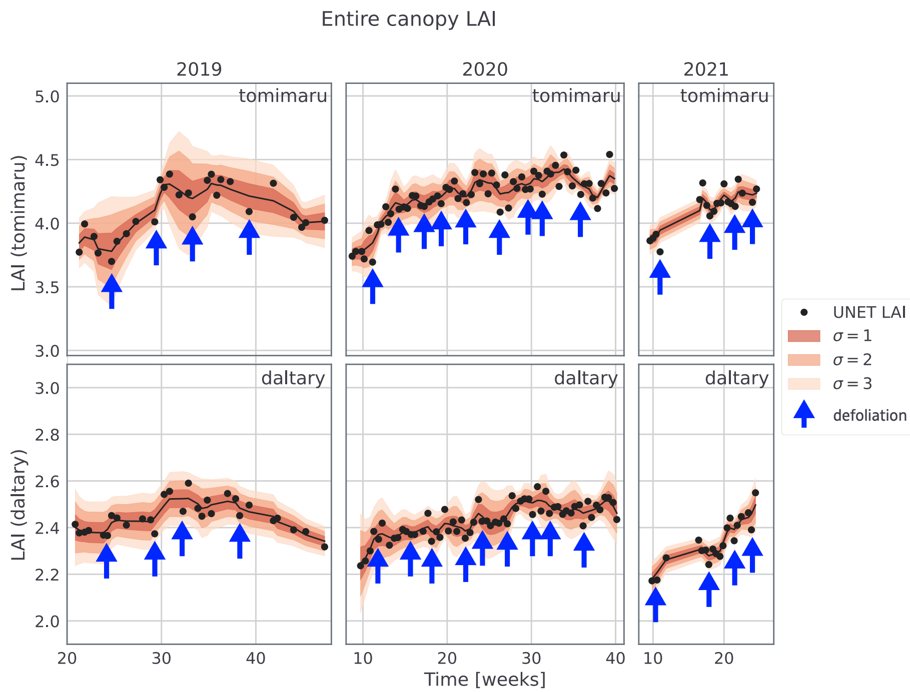

Figure 12: The **LAI of the entire plant canopy** computed with UNET semantic segmentation (black dots) with its time averaged spline interpolation (black line) for the period May 2019 to June 2021. The variability σ of the measurement data (2000 segmentation frames per observation) in shades of red. Defoliation instances are annotated with blue arrows.

The LAI estimations for the upper canopy 11 and the entire canopy 12 are not directly comparable because the entire canopy LAI estimation evaluates a much larger FOV, as presented in Figure 3, which leads to smaller variations within a single observation.

4. Discussion

The pairing of an automated quasi-static optical monitoring apparatus with modern image segmentation routines, such as those used in this study, is very accurate in determining variations in the LAI over time for complex vine structures, such as tomato plants. This approach is highly suited for industrial applications in greenhouse environments where plants are well aligned and managed. While the background and possible distractors can be sufficiently distinguished from the objects of interest and the training and reference loss is small enough (<0.2 and 0.1, respectively), a discussion about the accuracy of various segmentation algorithms appears to be unnecessary. This is because the general definition of the LAI is ambiguous and modern segmentation routines are sufficiently advanced.

Regarding the training and validation losses (as reported in section 3.2), it must be mentioned that complex plant images, as used in this study, are very difficult to annotate precisely. Further, the most challenging task is the differentiation between plants in the foreground (used to compute LAI) and plants in the background. This is because texture, lightning conditions and morphology are often similar. However, to estimate LAI, determining the exact segmentation morphology is not necessary, but needs to to be only precise enough to reflect the projected difference in leaf area. The variation of the segmentation area of the foreground leafs is smaller than three percent as presented in Figure 9. This is mainly because the foreground and background plants cannot always correctly be distinguished by the annotator. Therefore, the annotation quality (but not the segmentation quality) is highly dependent on the image detail, which is affected by the camera lens, sensor resolution, and lighting conditions.

In addition, one must consider that the baseline (LAI_{grower}) is fuzzy with high uncertainty. This is because computing the exact LAI for one tomato plant would require exact knowledge of the number of leaves and leaf area. In general, plant growth within greenhouse is mostly homogenous. However, it is possible that some plants growth diverges from that of the reference plants. This is the case in the for tomimaru plants grown during the year 2020 of our observation. Here, the reference plants used by the grower exhibited different LAI (lower until week 25, and higher from there on) than the rest of the greenhouse plants. This caused the grower to falsely estimate the amount of defoliation, resulting in an feedback loop of rising LAI for upper canopy. However, for all the observations of daltary and the remaining observations of tomimaru are in agreement with the upper canopy estimations of LAI_{grower}. Further, the LAI evolution, computed from evaluation the entire canopy is in agreement with expectations for tomato in an controlled greenhouse environment [13, 40]. Also, amplitude dips within the LAI evolution are in alignment with defoliation dates (Figure 12).

5. Conclusion

For non-quasi-stationary approaches (such as smartphone applications as well as other mobile solutions [24, 25, 16]), it is imperative to evaluate and understand plant morphology as well as the environment. The proposed approach does not require to perceive the exact plant and environment morphology, because, in the captured data, the plants and the environment are dimensionally and spatially homogeneous. This is ensured by the camera always traveling along the same path and covering the entire plant row.

The approach proposed in this study is more labor efficient than estimating the LAI manually but less efficient than a set of stationary cameras because, for the proposed approach, the robotic camera system must be maintained and observed during measurement. The authors believe that the degree of perception in this and more usual approaches (e.g., hand-held devices and sets of stationary cameras) could be improved by computing plant postures similarly to the human and hand posture estimations presented by Bazarevsky et. al. ([41, 42] and Liu et. al. [43]. From the plant posture, one could more precisely differentiate between foreground and background vines and also obtain higher spatial awareness.

6. Acknowledgements

We are grateful to Mr. Ohkura from Air Water and Prof. Ubukata from Takushoku University Hokkaido College for helpful discussions. We also thank Takeru Kanoh of Plant Data Co. Ltd. for the integrated camera system. This research was commissioned and supported by the National Institute of Information and Communications Technology (NICT), JAPAN, and the Adaptable and Seamless Technology Transfer Program through Target-driven R&D (A-STEP) of the Japan Science and Technology Agency (JST) (Grant Number: JPMJTM20A1). This work was also supported in part by JSPS KAKENHI (Grant Number No. 20K11968).

References

[1] A. AlKameli, M. Hammad, Automatic learning in agriculture: A survey, International Journal Of Computing and Digital System (2021).

[2] A. Kamilaris, F. X. Prenafeta-Boldú, Deep learning in agriculture: A survey, Computers and electronics in agriculture 147 (2018) 70–90.

[3] F. Rodríguez, M. Berenguel, J. L. Guzmán, A. Ramírez-Arias, Modeling and control of greenhouse crop growth (2015).

[4] X. Blasco, M. Martínez, J. Herrero, C. Ramos, J. Sanchis, Model-based predictive control of greenhouse climate for reducing energy and water consumption, Computers and Electronics in Agriculture 55 (1) (2007) 49–70. doi:https://doi.org/10.1016/j.compag.2006.12.001. URL https://www.sciencedirect.com/science/article/pii/S0168169906001165

[5] I. Herrmann, A. Pimstein, A. Karnieli, Y. Cohen, V. Alchanatis, D. Bonfil, Lai assessment of wheat and potato crops by ven s and sentinel-2 bands, Remote Sensing of Environment 115 (8) (2011) 2141–2151. doi:https://doi.org/10.1016/j.rse.2011.04.018. URL https://www.sciencedirect.com/science/article/pii/S0034425711001465

[6] R. E. Jongschaap, Run-time calibration of simulation models by integrating remote sensing estimates of leaf area index and canopy nitrogen, European Journal of Agronomy 24 (4) (2006) 316–324.

[7] Y. Fei, S. Jiulin, F. Hongliang, Y. Zuofang, Z. Jiahua, Z. Yunqiang, S. Kaishan, W. Zongming, H. Maogui, Comparison of different methods for corn lai estimation over northeastern china, International Journal of Applied Earth Observation and Geoinformation 18 (2012) 462–471. doi:https://doi.org/10.1016/j.jag.2011.09.004.
URL https://www.sciencedirect.com/science/article/pii/S0303243411001255

[8] H. Wang, J. Sánchez-Molina, M. Li, M. Berenguel, X. Yang, J. Bienvenido, Leaf area index estimation for a greenhouse transpiration model using external climate conditions based on genetics algorithms, back-propagation neural networks and nonlinear autoregressive exogenous models, Agricultural Water Management 183 (2017) 107–115, special Issue: Advances on ICTs for Water Management in Agriculture. doi:https://doi.org/10.1016/j.agwat.2016.11.021.
URL https://www.sciencedirect.com/science/article/pii/S0378377416304668

[9] R. Xu, J. Dai, W. Luo, X. Yin, Y. Li, X. Tai, L. Han, Y. Chen, L. Lin, G. Li, C. Zou, W. Du, M. Diao, A photothermal model of leaf area index for greenhouse crops, Agricultural and Forest Meteorology 150 (4) (2010) 541–552. doi:https://doi.org/10.1016/j.agrformet.2010.01.019.
URL https://www.sciencedirect.com/science/article/pii/S0168192310000420

[10] E. Heuvelink, P. Tijskens, M. Kang, Modelling product quality in horticulture: an overview (2003) 19–30.

[11] W. Shu, L. Wang, B. Liu, J. Liu, LAI estimation of cucumber crop based on improved fully convolutional network, CoRR abs/2104.07955 (2021). arXiv:2104.07955.
URL https://arxiv.org/abs/2104.07955

[12] C. Bacour, S. Jacquemoud, Y. Tourbier, M. Dechambre, J.-P. Frangi, Design and analysis of numerical experiments to compare four canopy reflectance models, Remote Sensing of Environment 79 (1) (2002) 72–83.

[13] W. J. Jo, J. H. Shin, Effect of leaf-area management on tomato plant growth in greenhouses, Horticulture, Environment, and Biotechnology 61 (6) (2020) 981–988.

[14] T. Popović, V. Maraš, S. Čakić, S. Šandi, S. Radonjić, K. Pavićević, Use of mobile applications in smart agriculture, PRACTICAL GUIDE FOR THE USE OF ICT IN AET 36.

[15] H. M. Easlon, A. J. Bloom, Easy leaf area: Automated digital image analysis for rapid and accurate measurement of leaf area, Applications in plant sciences 2 (7) (2014) 1400033.

[16] A. Patrignani, T. E. Ochsner, Canopeo: A powerful new tool for measuring fractional green canopy cover, Agronomy Journal 107 (6) (2015) 2312–2320.

[17] M. Mora, F. Avila, M. Carrasco-Benavides, G. Maldonado, J. Olguín-Cáceres, S. Fuentes, Automated computation of leaf area index from fruit trees using improved image processing algorithms applied to canopy cover digital photograpies, Computers and Electronics in Agriculture 123 (2016) 195–202.

[18] M. H. Radzali, N. A. M. Kamal, N. M. Diah, Measuring leaf area using otsu segmentation method (lamos), Indian Journal of Science and Technology 9 (48) (2016) 1–6.

[19] S. Fuentes, C. Poblete-Echeverría, S. Ortega-Farias, S. Tyerman, R. De Bei, Automated estimation of leaf area index from grapevine canopies using cover photography, video and computational analysis methods, Australian Journal of Grape and Wine Research 20 (3) (2014) 465–473.

[20] F. Orlando, E. Movedi, L. Paleari, C. Gilardelli, M. Foi, M. Dell'Oro, R. Confalonieri, Estimating leaf area index in tree species using the pock-

etlai smart app, Applied Vegetation Science 18 (4) (2015) 716–723. arXiv: https://onlinelibrary.wiley.com/doi/pdf/10.1111/avsc.12181, doi:https://doi.org/10.1111/avsc.12181.
URL https://onlinelibrary.wiley.com/doi/abs/10.1111/avsc.12181

[21] M. Campos-Taberner, F. J. García-Haro, . Moreno, M. A. Gilabert, S. Sánchez-Ruiz, B. Martínez, G. Camps-Valls, Mapping leaf area index with a smartphone and gaussian processes, IEEE Geoscience and Remote Sensing Letters 12 (12) (2015) 2501–2505. doi:10.1109/LGRS.2015.2488682.

[22] T. Lendzioch, J. Langhammer, M. Jenicek, Estimating snow depth and leaf area index based on uav digital photogrammetry, Sensors 19 (5) (2019) 1027.

[23] L. Comba, A. Biglia, D. Ricauda Aimonino, C. Tortia, E. Mania, S. Guidoni, P. Gay, Leaf area index evaluation in vineyards using 3d point clouds from uav imagery, Precision Agriculture 21 (4) (2020) 881–896.

[24] M. Campos-Taberner, F. J. García-Haro, R. Confalonieri, B. Martínez, . Moreno, S. Sánchez-Ruiz, M. A. Gilabert, F. Camacho, M. Boschetti, L. Busetto, Multitemporal monitoring of plant area index in the valencia rice district with pocketlai, Remote Sensing 8 (3) (2016). doi:10.3390/rs8030202.
URL https://www.mdpi.com/2072-4292/8/3/202

[25] R. De Bei, S. Fuentes, M. Gilliham, S. Tyerman, E. Edwards, N. Bianchini, J. Smith, C. Collins, Viticanopy: A free computer app to estimate canopy vigor and porosity for grapevine, Sensors 16 (4) (2016) 585.

[26] H. Fang, Y. Ye, W. Liu, S. Wei, L. Ma, Continuous estimation of canopy leaf area index (lai) and clumping index over broadleaf crop fields: An investigation of the pastis-57 instrument and smartphone

applications, Agricultural and Forest Meteorology 253-254 (2018) 48–61. doi:https://doi.org/10.1016/j.agrformet.2018.02.003. URL https://www.sciencedirect.com/science/article/pii/S0168192318300406

[27] H. Fang, W. Li, S. Wei, C. Jiang, Seasonal variation of leaf area index (lai) over paddy rice fields in ne china: Intercomparison of destructive sampling, lai-2200, digital hemispherical photography (dhp), and accupar methods, Agricultural and Forest Meteorology 198 (2014) 126–141.

[28] C. Payer, D. Štern, M. Feiner, H. Bischof, M. Urschler, Segmenting and tracking cell instances with cosine embeddings and recurrent hourglass networks, Medical image analysis 57 (2019) 106–119.

[29] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, S. S. Iyengar, A survey on deep learning: Algorithms, techniques, and applications, ACM Computing Surveys (CSUR) 51 (5) (2018) 1–36.

[30] S. Jung, S. Hwang, H. Shin, D. H. Shim, Perception, guidance, and navigation for indoor autonomous drone racing using deep learning, IEEE Robotics and Automation Letters 3 (3) (2018) 2539–2544.

[31] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F. E. Alsaadi, A survey of deep neural network architectures and their applications, Neurocomputing 234 (2017) 11–26.

[32] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, J. Garcia-Rodriguez, A review on deep learning techniques applied to semantic segmentation, arXiv preprint arXiv:1704.06857 (2017).

[33] M. L. Tenzer, N. C. Clifford, A digital green thumb: Neural networks to monitor hydroponic plant growth (2020) 1–6.

[34] Y. Liu, D. Minh Nguyen, N. Deligiannis, W. Ding, A. Munteanu, Hourglass-shapenetwork based semantic segmentation for high resolution aerial imagery, Remote Sensing 9 (6) (2017) 522.

[35] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation (2015) 234–241.

[36] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library (2019) 8024–8035.
URL http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

[37] K. Yamaguchi, M. Korobov, et al., psd-tools: Open source psd tools for Python (2019–).
URL https://psd-tools.readthedocs.io

[38] E. Jones, T. Oliphant, P. Peterson, et al., SciPy: Open source scientific tools for Python (2001–).
URL http://www.scipy.org/

[39] S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, scikit-image: image processing in python, PeerJ 2 (2014) e453.

[40] T. Saito, Y. Mochizuki, Y. Kawasaki, A. Ohyama, T. Higashide, Estimation of leaf area and light-use efficiency by non-destructive measurements for growth modeling and recommended leaf area index in greenhouse tomatoes, The Horticulture Journal 89 (4) (2020) 445–453. doi:10.2503/hortj.UTD-171.

[41] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, M. Grundmann, Blazepose: On-device real-time body pose tracking, arXiv preprint arXiv:2006.10204 (2020).

[42] V. Bazarevsky, F. Zhang, On-device, real-time hand tracking with mediapipe, Google AI Blog (2019).

[43] L. Liu, J. Xing, H. Ai, X. Ruan, Hand posture recognition using finger geometric feature (2012) 565–568.