UNIVERSITY OF CALIFORNIA
RIVERSIDE

Vision Based Detection of Xylem Water Expression

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Electrical Engineering

by

Joshua J. Chen

June 2022

Thesis Committee:

Dr. Konstantinos Karydis, Chairperson
Dr. Amit K. Roy-Chowdhury
Dr. Stefano Carpin

The Thesis of Joshua J. Chen is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

The author gratefully acknowledges the support of USDA-NIFA under grant # 2021-67022-33453, and a UC MRPI Award. Any opinions, findings, and conclusions expressed in this material are those of the author and do not necessarily reflect the views of the funding agencies. The author would also like to thank the University of California, Riverside Agricultural Operations for permitting access to their orchards to conduct field experiments and data collection.

I am grateful for:

Alice, my girlfriend, whose endearing company and warm embrace I cannot do without.

Terry, my childhood friend, whose generosity and mutual cooking interests provided me a place to delve into my hobby.

Adrian, Michael, and Mikey. My undergraduate ME cohort friends who were always by my side and sharing relevant class memes.

Amel and Merrick. My graduate friends who imparted great advice and provided guidance.

Dr. Karydis, my advisor, for providing me an opportunity to expand and develop my interests and understanding in automation.

ABSTRACT OF THE THESIS

Vision Based Detection of Xylem Water Expression

by

Joshua J. Chen

Master of Science, Graduate Program in Electrical Engineering
University of California, Riverside, June 2022
Dr. Konstantinos Karydis, Chairperson

Stem water potential (SWP) measurements are used in agriculture to determine the water stress in crops and the optimal irrigation schedule. The current widely used method to take such measurements is the Scholander pressure chamber due to its simplicity, but it is labor intensive and presents potential harm to the operator if not prepared or used properly. Automating the process of taking SWP measurements with a computer vision control system can help increase efficiency, precision, and accuracy over the current methods, additionally ensuring operator safety as it is unnecessary for the operator to maintain close proximity to the chamber. This research aims to pioneer an autonomous computer vision control system as a means to aid growers and function on the existing platform, with the intent of eventually being widely adopted in precision agriculture.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Agriculture has long been a primarily manual sector, with many of its facets featuring little to no automation. However within the most recent decades, efforts to improve agriculture through the incorporation of technologies from other sectors [1] has given rise to a more automated, precision agriculture. At the core of our research is a specific aspect of precision agriculture, precision irrigation, which intends to optimize irrigation patterns across various crops to generate greater yield and to reduce environmental impacts related to water usage [2]. To optimize irrigation patterns, agronomists rely on a metric called water potential to determine a crop's water stress level [3, 4]. More specifically, agronomists will measure water potential through leaf stem water potential (SWP) analysis with a pressure chamber. The returned pressure reading at water expression is then used as a proxy for water stress level in the crop since the pressure is an indicator of water potential [5, 6]. This measurement then informs agronomists of the optimal irrigation schedule.

The process of conducting SWP analysis currently requires an agronomist to visit an orchard and conduct measurements in-field. It is proper protocol to first bag the leaf sample with a reflective foil bag for at least ten minutes to mitigate transpiration during the measurement process [7]. After which, the leaf sample is excised cleanly at the stem and placed into a pressure chamber with the xylem exposed to the atmosphere. As the chamber pressurizes, the agronomist will manually observe water expression at the excised xylem through a magnifying glass and note the pressure at which a water drop forms. This entire process is labor intensive, time consuming, and presents potential bodily harm[1] to the pressure chamber operator through turning the excised xylem into a projectile or expelling particles through the xylem into the operator's eyes [8]. Given these grievances, it is often that only a handful of measurements are used to quantify the state of an entire orchard, resulting in potentially misidentifying the health of an orchard. This inefficiency is what our work intends to resolve, through automating the SWP analysis process with a computer vision control system.

1.2 Overview

Automating the SWP analysis process with a computer vision control system to detect xylem water expression will increase the safety, precision, accuracy, and efficiency over the current conventional method. This is due to the automated system not requiring the operator to remain in close proximity to the chamber nor for the operator to input any manual labor. Additionally, a computer vision control system has a more precise

[1] Chamber Safety: `https://www.pmsinstrument.com/maintenance/safety/`

definition of water expression, whereas different people will have differing definitions of water expression, so manual measurements may differ among separate agronomists. As such, this work aims to develop a computer-vision based approach to detect water expression to help automate part of the SWP analysis process.

We develop and test a learning-based computer vision architecture based on the YOLOv5 object detection network. We evaluate five distinct variants of the network to better understand the effect of various parameters and identify the setup that yields the best results. The validity of our computer vision control system is shown through the results yielded by two specific variants; the networks yielded mean average precision (mAP) scores of 99.5% with nearly ideal precision and recall values on the validation set. They also managed to classify the test set images with a success rate of 99.16%, where the test set images are unique to both the train and validation set images, and hence not previously seen by the network.

1.3 Organization

This thesis follows a chronological structure that resembles the workflow of the research presented. First, the conceptual background and foundation for the work presented herein are explained. Next, the conceptualization phase is explored, with explanations of design decisions and details regarding hardware development. Afterwards, details regarding collecting and processing the image dataset are presented, along with an overview of the computer vision architecture and training process. Finally, the system validation and results are discussed, along with plans for future developments.

Chapter 2

Conceptual Background

This chapter presents the conceptual foundations for our research and explains the relevance of past and alternative research endeavors in obtaining stem water potential measurements. The indications of stem water potential are explained, followed by an exposition of the pressure chamber method, alternative paths, and related computer vision topics.

2.1 Stem Water Potential

As mentioned earlier, stem water potential (SWP) is a metric agronomists rely on for scheduling and optimizing irrigation patterns [3, 4]. It is a direct indicator of crop health since water potential refers to the propensity for water to move along a gradient from the soil, up the plant, and out through transpiration. Greater water potential translates to higher water stress, meaning the plant needs hydration. There are multiple currently or previously employed methods to measure SWP [9], with the pressure chamber being one of the most popular and a benchmark comparison for other methods [10, 11].

Figure 2.1: Insertion of leaf sample into reflective foil bag.

2.2 Scholander's Pressure Chamber

The Scholander pressure chamber was first introduced in the 1960s as a means of measuring SWP since previous methods have proved ineffective [12]. The pressure chamber method involves inserting a leaf sample into an empty chamber, and having the stem of the leaf protrude out of the chamber through a single interface, flanked by a rubber compression gland or o-ring on all sides to grip and stabilize the leaf during water potential measurement.

As a general practice now, leaves are covered with a reflective foil bag (Figure 2.1) before being excised and during the measurement process in the chamber to mitigate water loss through transpiration and to prevent faulty SWP measurements. The difference of

water potential, or hydrostatic pressure needed to reach the wet point, can vary 2-3 bar between bare leaves and aluminum foil covered leaves [7] despite inserting the samples into a pressure chamber within 30 seconds after excision. It is also now frequent to perform these measurements during midday to early afternoon where the photosynthesis rate is highest to minimize measurement variation and to obtain the most accurate water potential measurements [13].

The premise for the pressure chamber method is that a sufficiently high hydrostatic pressure applied to the leaf counteracts the cells' osmotic potential, and forces the intracellular fluids outward towards atmospheric pressure [14, 15], in this case the excised xylem shown in Figure 2.2. According to Boyer, this relationship can be summarized as

$$\Psi_w = P + \Psi_s \tag{2.1}$$

where Ψ_w represents the osmotic potential of the plant cells, P represents the hydrostatic pressure within the xylem, and Ψ_s represents the osmotic potential of solutes in the xylem sap. This osmotic potential, or water tension, within the plant is caused by a gradient between the soil moisture and leaf transpiration rate [3]. It is important to note that dryer soil results in higher water tension within plants, which in turn affects the plant's carbon sequestration capabilities and growth rate [16]; therefore, stem water potential measurements are greatly significant for agronomists, and our work aims to ease the process of obtaining these measurements through an automated pressure chamber method.

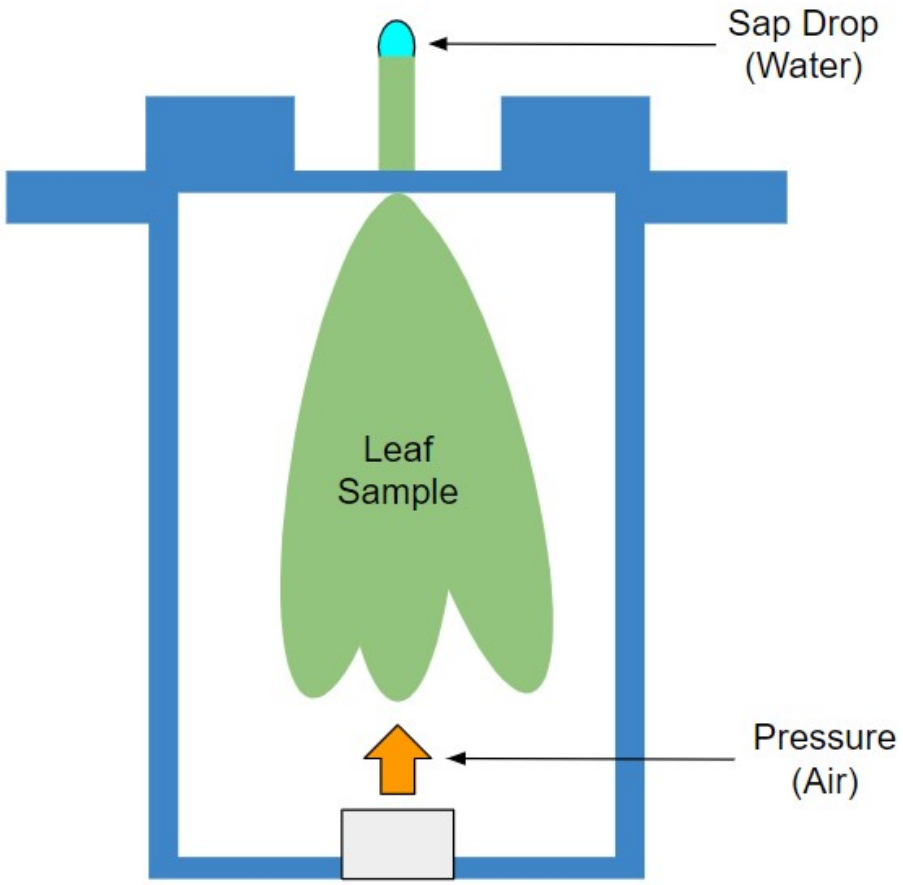

Figure 2.2: Working concept of Scholander's pressure chamber.

2.3 Alternate Methods

Since optimizing irrigation patterns relies heavily on water potential measurements, there have been some recent alternative endeavors on obtaining these metrics. The two newly proposed methods, in addition to being non-destructive, offer scalability over the current pressure chamber method as they can be used for fast bulk assessment of crops. However, both methods are extremely sensitive to weather and light conditions due to the nature of their operating principles.

Zhao et al. proposed a new method based on employing multi-spectral cameras on small UAVs to perform high resolution multi-spectral imaging of crops using the canopy Normalized Difference Vegetation Index (NDVI). The multi-spectral images would then be used for water potential prediction. However, it was mentioned that the data collected during the same day from different flights exhibited high variability due to solar motion [17].

Another newly proposed method, by Vila et al., relies on the use of remote sensors and spectral reflectance of the plant to measure water potential. Although initially promising, the results yielded low correlation and thus concluded the spectral reflectance method cannot serve as a viable replacement for the pressure chamber method [18]. These spectral based methods may not yet be fully developed and may not yet be ready for measuring water potential, but they have provided us insight on our own xylem imaging task for our vision model.

2.4 Computer Vision Related Works

Detecting water expression at the excised xylem requires the computer vision system to be capable of recognizing the subtle feature variance present on the xylem cross section. As presented later in Figure 4.2, during water expression, the sharpness, hue, and saturation of the pith change due to moisture. Additionally, the light refraction characteristics change due to the presence of the sap drop. All of these features can be detected through a convolutional neural network (CNN) and object detector. Patel et al. demonstrated that wetness or water drops on the surface of plants can be detected through the use of a deep CNN with RGB colored images to a high degree of accuracy and reliability [19]. Fuentes et

al. presented color invariant features of tomato crop diseases that can be detected through a CNN [20]. Building upon these findings, we opted to approach our computer vision system with a CNN-based object detector, and relevant training image augmentations (applied from third version onward) that help emphasize features of interest.

Chapter 3

Conceptualization & Hardware Development

This chapter explores the conceptualization phase of our research, where details and design considerations are explained in depth. Crucial points of focus for this phase include the selection processes for the pressure chamber and cameras, as well as camera mount design.

3.1 Pressure Chamber Choice

Our main considerations when choosing a pressure chamber were the portability and convenience of operating the device. It was agreed that a portable chamber that can be operated with minimal components and one that can be used for rapid prototyping of the computer vision control system, such as data collection, would be the ideal design for

Figure 3.1: PMS Instruments Pump-Up Pressure Chamber (Brand: PMS Instruments).

the early phases of this research. With these constraints considered, the choice for our pressure chamber was the PMS Instruments[1] pump-up pressure chamber in Figure 3.1. This chamber's conservative size, portability, and ease of use allowed us to collect image data on avocado crops easily both in-lab and in-field. It requires no electrical power or compressed air source, as the chamber is pressurized manually by an operator pumping the piston repeatedly. These attributes proved helpful during rapid prototyping as well when different camera mounts were being tested on the chamber body.

[1]PMS Instruments: `https://www.pmsinstrument.com/`

One may raise the concern of using a small pressure chamber to conduct our experiments since many crops' leaves and stems do not necessarily fit nicely into the internal space, let alone the reflective bags that come with the pressure chamber. To this we emphasize the focus on using only avocado crops for our experiments due to keeping the debugging process consistent. Additionally, many of the smaller avocado leaves from our plants fit into the chamber and the typical pressures seen by avocado (within 20 bar) [21] are within the safe operating range of the pump-up chamber, making avocado trees an ideal crop to test on for this research.

During the later stages of this research and subsequently our future works, another pressure chamber model was selected for validating the combined computer vision control system and physical system. This model is not a manual pump-up design, but rather a suitcase type design that requires a compressed air source to pressurize the embedded chamber. The advantages of this model over the pump-up include the higher pressure tolerance and lack of movement, which will help the computer vision model immensely when detecting water expression.

3.2 Camera Selection

The camera selection process involved determining a suitable camera for our xylem imaging task. Given the diametral magnitude of the excised xylem to be approximately only 2 mm, the constraints specified cameras with close-up magnification and focal capabilities, but with less magnifying strength than microscope cameras. Logically, we determined the

Figure 3.2: Endoscope Camera used for data acquisition with opaque tape modification (Brand: Suear).

ideal candidates to have a point-blank minimum object distance [22] (MOD), that is, the minimum distance between the object and camera for focus to be possible. The ideal MOD we determined is within 7 cm as this precludes much of the unnecessary background in the camera feed and results in a higher resolution of the xylem's cross section.

The first idea that materialized was an endoscope camera, or "endocam" for short. This camera has a minimum object distance of 2-3 cm, which satisfies our close-up MOD requirement and also happens to be the focal range. The endoscope camera selected for our work is shown in Figure 3.2, and features a wireless design. Although useful for the early phase of our imaging task with the setup in Figure 3.3, it features no method to communicate with a computer in real time to stream video feed, so it proved to be practical for data collection only.

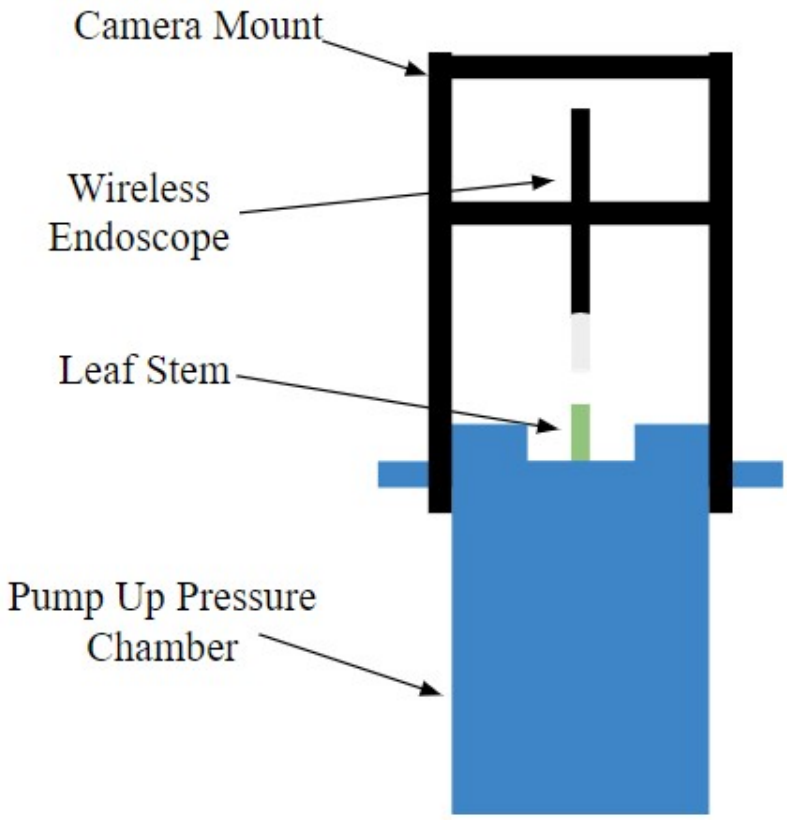

Figure 3.3: Endocam setup for data acquisition.

A more practical choice was reconsidered as we used the endocam for the first half of data collection. Since we need the camera to communicate with a computer to stream video data, we opted for the HQ Picamera, which is compatible with the Raspberry Pi single board computer. The HQ Picamera in Figure 3.4 features a high resolution of 12.3 megapixels and a mount to attach C and CS mount lenses. For our work, we mounted a 16 mm telephoto lens to the HQ Picamera with magnification abilities to observe the viability of two setups: Reverse lens macrophotography[2] and Offset lens.

[2]Reverse Lens Macro Technique: https://dev.webonomic.nl/macro-photography-with-the-raspberry-pi-hq-camera-and-reversing-the-lens

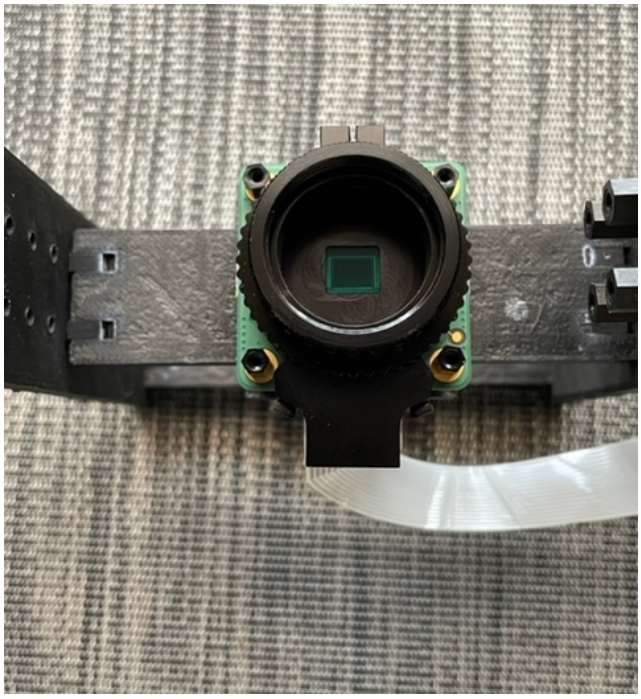

Figure 3.4: High Quality Picamera (Brand: Raspberry Pi Foundation).

For the reverse lens macro setup in Figure 3.5, we attached the HQ Picamera to a reverse mounted 16 mm telephoto lens via adapters, and connect it all to a Raspberry Pi 3B+ single board computer. This setup mounts the telephoto lens in the reverse orientation because telescopic lenses offer great magnification abilities when mounted in reverse [23]. When testing this setup, we discovered the magnification was too great and the depth of field was too shallow, so it was difficult to focus on the entire xylem because the xylem exceeded the bounds of the frame and focusing the camera with an extremely shallow depth of field proved difficult and impractical. Additionally to correct the shallow depth of field issue, the lens aperture was increased, but this presented a lighting issue where the feed was underexposed.

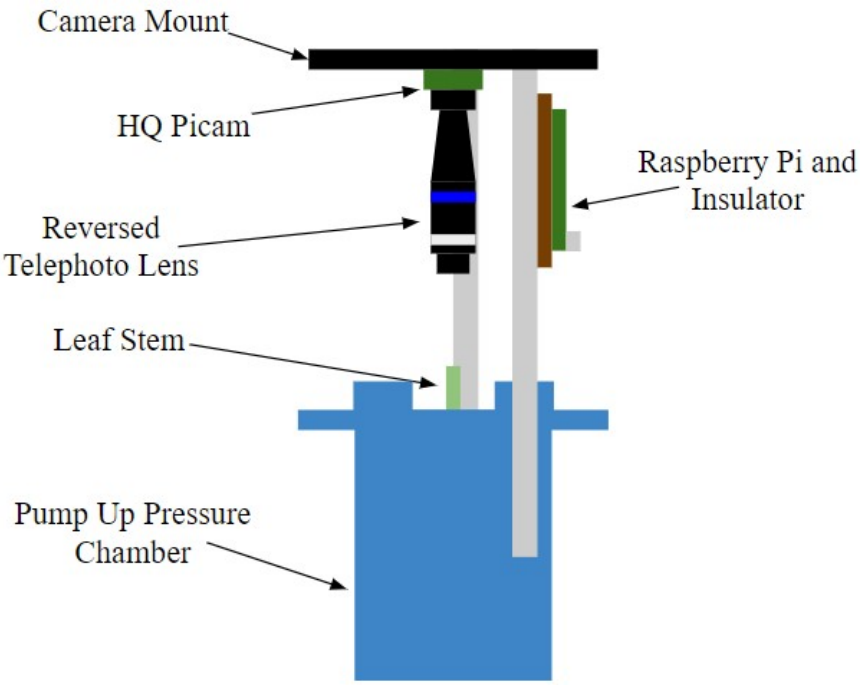

Figure 3.5: Reverse Lens Macro Setup with HQ Picamera.

More success was seen with the offset lens setup shown in Figure 3.6, where the telephoto lens was mounted in the nominal orientation, but with an offset adapter between the HQ Picamera and the lens itself. The effects of this technique include decreasing the MOD of the camera and increasing magnification. This allowed us to mount the camera closer to the stem and preclude most of the unnecessary background from the video frames. An indispensable advantage the offset camera setup has over the reverse lens setup is the ability to adjust focus easily. What resulted were highly defined and magnified images of the xylem, perfect for training our computer vision model in the subsequent phases of the project.

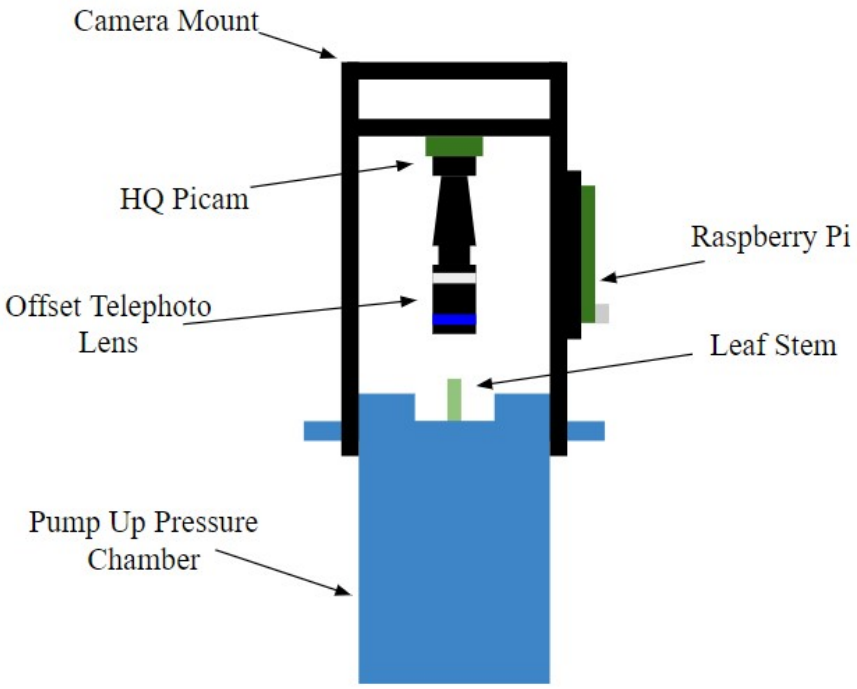

Figure 3.6: Final HQ Picamera setup using offset lens technique.

3.3 Camera Mount

Since both the endocam and HQ Picamera offset configurations offer excellent xylem imaging, it was decided to use both of these setups for data acquisition. The camera mount designed in this work considers mounting both setups on the same platform, but separately, to reduce fabrication material usage. The Oynx 3D printed mount assembly features two links to grip onto the removable chamber lid, two platforms screwed onto the links to provide lateral stability, and a transitional fit interface to securely mount the endocam. In addition to the endocam interface, the bottom platform component also features four screw mounts to secure the HQ Picamera as shown in Figure 3.7.

Figure 3.7: HQ Picamera offset lens configuration.

In regard to the endocam setup specifically, it may seem unstable to mount the endocam through only one platform. Despite this, the endocam does not exhibit any oscillatory motion and is practically rigid due to the minimal clearance in the platform interface. Although the mount is stable, due to the nature of using a manual pump-up pressure chamber to collect the data, vibrational noise can not be zero. Regardless, the preprocessing

steps outlined in Sec. 4.3 delete all unusable frames from the video data, and retain the clear frames where the xylem is in focus.

An additional modification made to the endocam is the addition of opaque tape around the camera end of the endocam, shown later in Figure 4.1. This serves as an improvised aperture control solution since the device features no built-in aperture control and the LEDs are overexposing the video feed. Collecting usable data where water expression can be easily observed requires the xylem cross section to be completely visible with no obstructions and appropriate exposure.

Chapter 4

Image Dataset Collection

The image dataset of this research consists of approximately 4100 images of dry and wet xylem cross sections of avocado crops taken from 11 separate videos recorded during stem water potential measurements with the PMS Instruments manual pump-up pressure chamber. All videos were recorded through two different camera setups: Endocam and HQ Picamera. Both camera setups provided excellent magnification of the xylem cross sections from a top-down perspective and presented clear states for the computer vision model to learn.

4.1 Collection Procedures

Dataset collection was rooted in the use of macrophotography capable cameras or macrophotography techniques to produce magnified images of the xylem cross section, while keeping the camera's minimum object distance within 7cm to constrain the camera mount to a conservative size. The videos were recorded with both the endocam and HQ

Figure 4.1: PMS pump-up pressure chamber retrofitted with the 3D printed endocam setup, enabling real-time video feed of the xylem during SWP measurements.

Picamera on a 3D printed mount that fits both setups separately, and could be secured to the top of the pump-up pressure chamber. Figure 4.1 demonstrates this with the endocam setup used during outdoor SWP data collection.

The process of taking SWP measurements and collecting video data required following the procedures outlined by Meron et al [7]. The avocado leaves were bagged with a reflective foil bag for at least ten minutes before excision. Once excised, each leaf was

then immediately transferred to the pressure chamber for measurement to mitigate water loss since the ambient conditions that day were hot and dry. Video recording was initiated only after the pressure gauge displayed at least 5 bar. This was done to decrease video file size, and subsequently make the preprocessing phase more convenient. The value of 5 bar was chosen specifically because avocado SWP measurements typically yield values around 12 bar [21]. Table 4.1 presents the SWP pressure measurements conducted during one of the data collection periods with ten leaves cut by one of two methods. As expected from Sharon et al., the average SWP pressure measurements were approximately 12 bar.

Table 4.1: Stem Water Potential Measurements

Cut Method	# of Leaves	Avg. Pressure (Bar)	Std. Dev (Bar)
Manual	5	11.34	0.96
End-effector	5	10.84	1.14

4.2 Notable Features

Upon reviewing the collected footage, we noticed that most stems exhibit particular characteristics during the onset of water expression. Most notably, the xylem's central region contains a profoundly white pith when dry, which desaturates into the surrounding green hues as moisture increases at the stem's excision. Additionally, either a sap drop or bubbles are visible at full water expression. These notable features can be extracted and learned by the computer vision model in differentiating dry and wet xylem states. Figure 4.2 displays these features in transition from dry to wet in juxtaposed frames from the endocam setup.

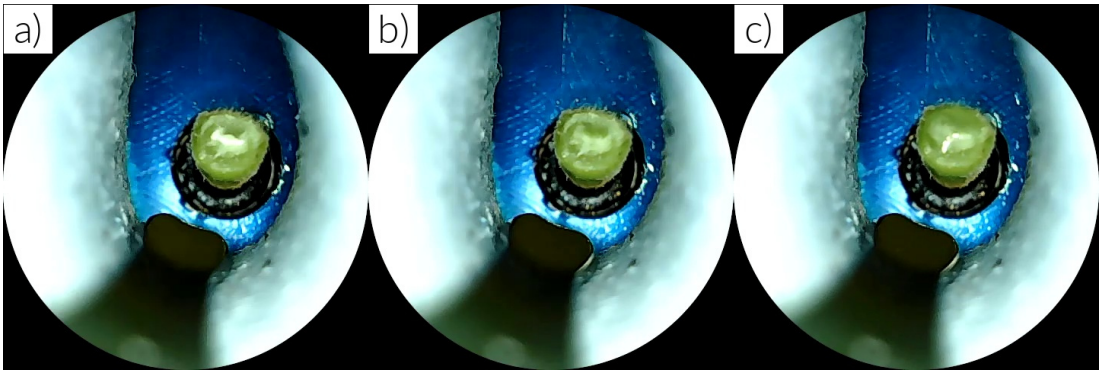

Figure 4.2: a) Xylem Dry State b) Xylem Transitional State c) Xylem Wet State.

4.3 Preprocessing

Preprocessing the images and videos in the dataset was necessary given the nature of operating the pump-up pressure chamber. Although the camera mount was designed to stabilize the camera assemblies over the chamber interface, vibrational noise was still introduced into the videos. The noise in the videos occurred primarily at the end of each upstroke and downstroke, when the chamber's internal pneumatic components would collide at their endpoints. This was inevitable for the particular pressure chamber used throughout the majority of this research, but the noise was corrected through frame omission.

Since the dataset created in this work relies on images and not videos, the solution to removing noise was omitting the frames where movement was present and/or the xylem was not in focus. This was performed by using a desktop annotation tool[1] to annotate or leave void some frames within a video, after the frames had been extracted to a working directory as images. Once all images have been labeled with a binary class label (0-Dry; 1-Wet) or left blank, a Python script developed in this work was executed to delete zero-

[1] Annotation Tool: `https://github.com/ManivannanMurugavel/Yolo-Annotation-Tool-New-`

area annotation files and their corresponding frames from the working directory. The result is a folder of usable images and their corresponding annotation files that are ready to be partitioned into training, validation, and testing sets.

Within the dataset, only specific images along the trimmed timeline for each video are used. The images used are from only the clearly dry and clearly wet regimes, no transitional (between dry and wet) regime images are used. This video trimming was performed to enforce precision regarding the definition of the wet point. With the conventional pressure chamber method, different agronomists may have slightly differing definitions on what constitutes as the wet point in SWP analysis, so imprecision is an issue. By omitting transitional regime frames, we are standardizing and keeping the wet point definition consistent by defining the wet point as when water expression yields a full, distinctive water drop (or mass of bubbles) on the excised xylem surface.

Chapter 5

Computer Vision Architecture & Training

This chapter presents the conceptual background for the software aspect of our work, along with details regarding the choice of our computer vision model and the training procedures that follow. Our object detector was trained using Google Colab Tesla T4 GPUs to expedite the training process. In total, five distinctive models were trained with varying parameters and training images, with versions 3 and onward being trained with image augmentations.

5.1 Model Choice

Based off of the related works presented in Sec. 2.3 and Sec. 2.4, we decided to approach the computer vision task with a convolutional neural network (CNN) based object detector. A CNN alone would not have been appropriate for the task since we are not trying

to classify an entire image as one of two binary classes, but rather to detect a specific region of each image (the xylem) and recognize it as dry or wet.

The CNN based object detector model we chose to use for our computer vision architecture is the YOLOv5 network[1], developed by Glenn Jocher and the Ultralytics[2] team. It is based on the previous network version YOLOv4, but boasts faster inference speed with similar accuracy metrics. So in addition to the high classification accuracy offered by YOLOv4 [24], YOLOv5 offers great potential in real-time applications. Furthermore, the YOLOv5 implementation is written in PyTorch[3] [25], an open-source machine learning framework which eases development and deployment.

5.2 Training Partition

Before training can begin, the image dataset was partitioned into training, validation, and testing sets. The training and validation sets were based off of the same xylem images, while the testing set was based off of a separate unique set of xylem images to simulate live inference. The ratio of the partition is approximately 80% (3232) training images and 20% (809) validation images, along with 119 separate images for the testing set. These metrics do not include image augmentations seen in the training of versions 3-5 of our model. As for some training parameters, all models were trained with batch sizes of 16 images and at least 25 epochs.

[1]YOLOv5: `https://github.com/ultralytics/yolov5`
[2]Ultralytics: `https://ultralytics.com/yolov5`
[3]PyTorch Open Source Machine Learning Framework: `https://pytorch.org/`

5.3 Augmentations

Training the final three versions of our computer vision model differed from the previous versions with the addition of applying image augmentations to the training data. As will be discussed in greater detail later in Sec. 6.2, the results yielded by the first two versions of the model were subpar, and image augmentations were applied to the training of V3, V4, and V5 to improve the performance over V1 and V2.

Augmenting training images carries many benefits, such as building a more robust model and increasing model performance. This occurs since augmenting training images increases the dataset size and diversity [26], leading to a more well-rounded model that does not overfit to a particular set of situations. Both V1 and V2 exhibited this issue since they could classify correctly only the water expression cases that they were trained on. For cases that involved excessive bubbling or extremely subtle water expression that were not present in the initial training sets, these two versions had immense difficulty classifying such cases in the testing sets. However with the training of the final versions, these issues were surmounted with the addition of more instances of unseen cases and image augmentations as seen in the batch of training images in Figure 5.1. Additionally, mosaic augmentation was applied to help the models recognize the xylem in potentially busier backgrounds [27] from the endocam images.

To apply image augmentations to the training data and ensure compatible formatting with the PyTorch framework, we used the recently incorporated Albumentations[4] library to augment the training data seamlessly for YOLOv5. The specific image augmen-

[4]Albumentations: `https://albumentations.ai/`

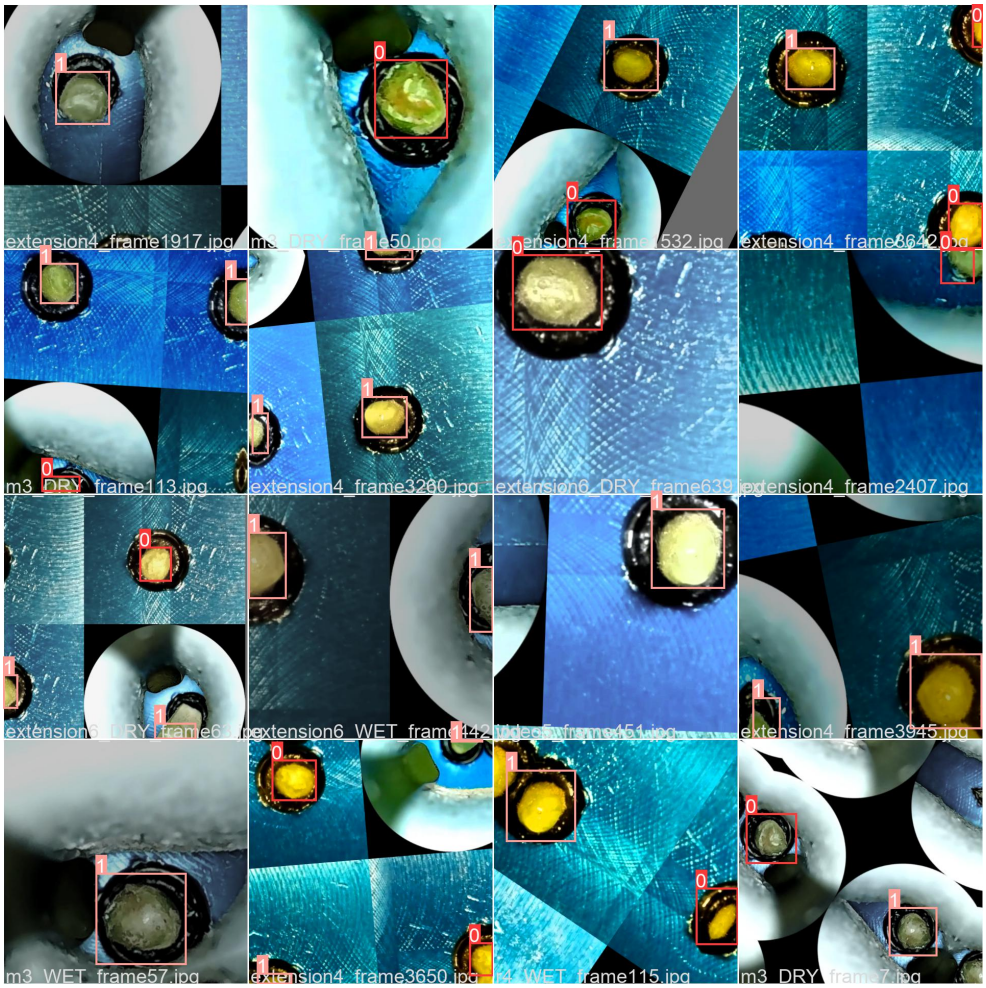

Figure 5.1: A set of training batch images with augmentations for V3.

tations applied are listed below with their details and justifications of use, with Gaussian Blur and HSV being exclusive to V4 and V5 respectively:

- Rotation: The input image is rotated by a random angle between -90 and 90 degrees with a probability of 0.5.
- Flipping (Horizontal and Vertical): The input image is flipped either horizontally or vertically with a probably of 0.5 for each augmentation.

- CLAHE (Contrast Limited Adaptive Histogram Equalization): The input image is enhanced to mitigate fogginess [28] with a probability of 0.1. (0.2 for V5)

- Random Brightness Contrast: The input image has its brightness and contrast randomly adjusted. For our training, the default settings for brightness were changed so this augmentation would not decrease image brightness, but rather only increase image brightness with a probability of 0.15. (0.2 for V5)

- Random Gamma: The input image is subjected to midtone contrast and brightness adjustments randomly with a probability of 0.1.

- Sharpen: The input image is sharpened and the result is superimposed on the original image with low opacity. This augmentation occurs with a probability of 0.1. (0.2 for V5)

- Gaussian Blur: The input image is subjected to a Gaussian mask of size 3x3 to 9x9 randomly with a probability of 0.15. This augmentation is applied to only V4.

- HueSaturationValue (HSV): The input image has its hue, saturation, and luminance randomly shifted with a probability of 0.15. This augmentation is applied to only V5.

The typical spatial-level transforms like rotation and flipping are applied because the orientation and appearance of every xylem is unique. Applying rotation and/or flipping to some training instances simulates "new" xylem profiles that may be observed in a realistic setting. The remaining augmentations are pixel-level transforms, and they are applied to account for a more diverse range of lighting conditions that may occur during live testing.

Chapter 6

System Validation

This chapter presents the test results yielded by our computer vision model and discussion of those results. In total, there are five distinctive versions of the computer vision model, with each version generally improving upon the capabilities of the previous, except for V4. The first two versions are not trained with augmentations, but the final three versions are trained with several different image augmentations. Since the final versions use training data augmentations as a regularization technique, they are more resilient to overfitting than the first two versions. The accuracy-loss quality metrics confirm this, thus validating our computer vision model and its potential in automating the SWP analysis process.

6.1 Testing

Our computer vision model was tested and validated through running inference on Google Colab and its GPUs. The testing set consists of 119 unique xylem images not

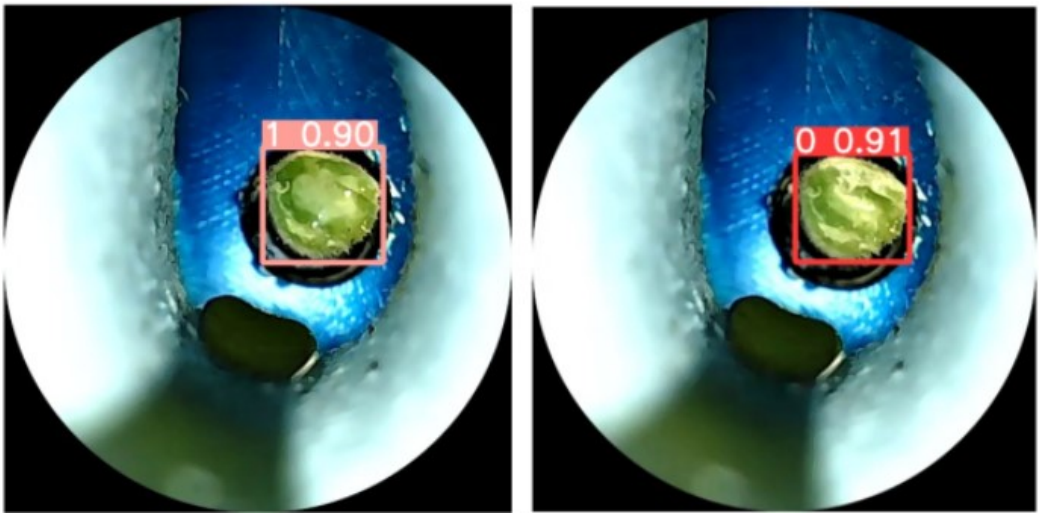

Figure 6.1: Correctly classified ambiguous case by V3.

seen by the model during training. This was done to verify the model is robust enough for actual live inference later on in this research endeavor. The first of the five models trained, V1, was trained on only approximately 2400 images since this was the initial dataset in the earlier phases of our work. The subsequent V2 was trained on approximately 2900 images, capturing more water expression cases but still featuring no augmentations. The final versions were trained on approximately 3200 images, capturing even more water expression cases including the ambiguous case in Figure 6.1, and featuring a number of augmentations.

6.2 Results

6.2.1 Version 1

The results of V1 were subpar at best, scoring well on the validation set but classifying the majority of the test set incorrectly. Initially scoring a mAP score of 99.5%

on the validation set, V1 seemed to possess great potential, but it was extremely overfitted to only the cases it saw during training, so it was completely lost when classifying test images. This is explained by the lack of using regularization techniques such as augmentations, and the limited exposure to general water expression cases during training. The immediate and obvious solution to improve V1 was to train it with more data.

6.2.2 Version 2

To remedy the shortcomings of V1, V2 was trained for 50 epochs on more cases of water expression with several more different xylems to combat overfitting. Much like V1, V2 also scored highly on the validation set, with a mAP score of 99.5%, but it also was still overfitted as it was incorrectly classifying particular cases it was not trained on, such as the ambiguous case in Figure 6.1 where the dry and wet states feature only a subtle

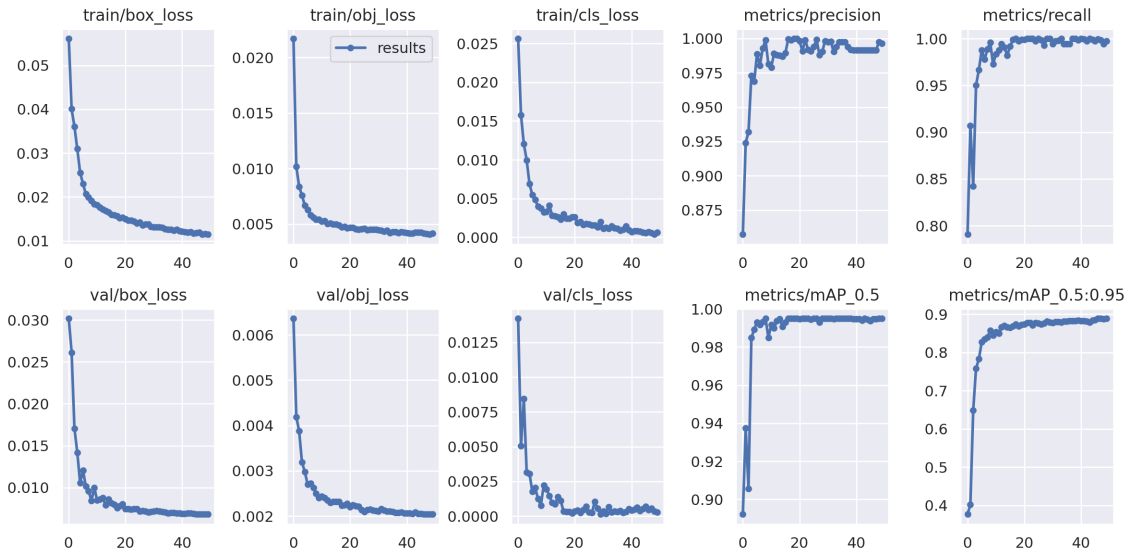

Figure 6.2: V2 Quality Metrics.

difference between them. For V2, the training loss converges at approximately 40 epochs (Figure 6.2), and this is largely the same with the validation loss. However as mentioned, V2 is still overfitted because when considering validation classification loss, the values do not converge cleanly, as the loss starts to increase slightly above 40 epochs. Despite this, V2 did experience an overall performance increase from V1 because it was able to correctly classify some of the cases that stumped V1.

6.2.3 Version 3

V3 yielded wonderful and accurate classification results, thus becoming the first robust model we have trained as of yet. The training and validation loss for V3 start to converge after 20 epochs (Figure 6.3), albeit with some minor fluctuations in the training loss due to using image augmentations as a regularization technique. In comparison to V2,

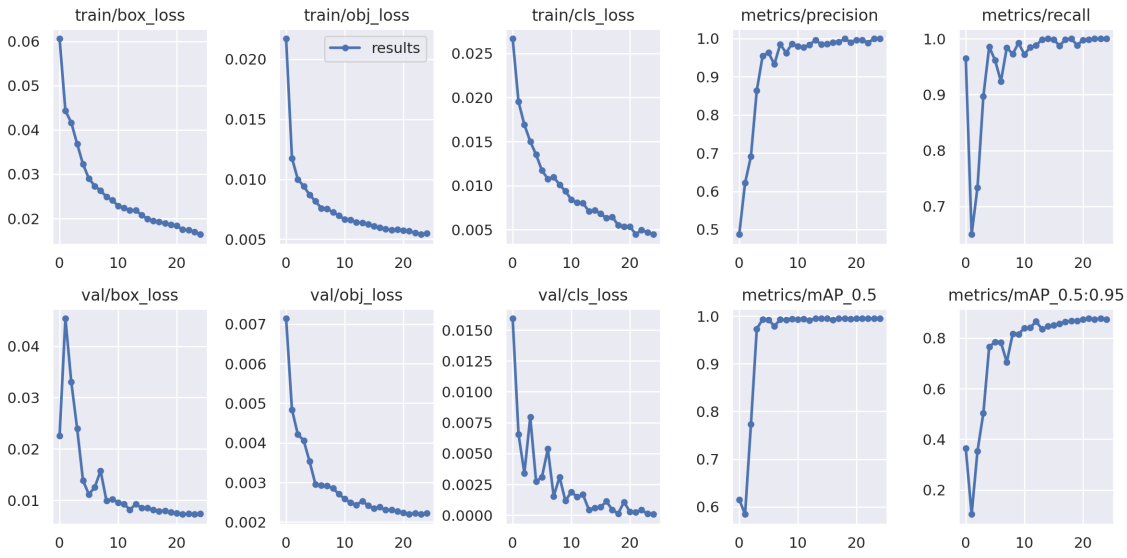

Figure 6.3: V3 Quality Metrics.

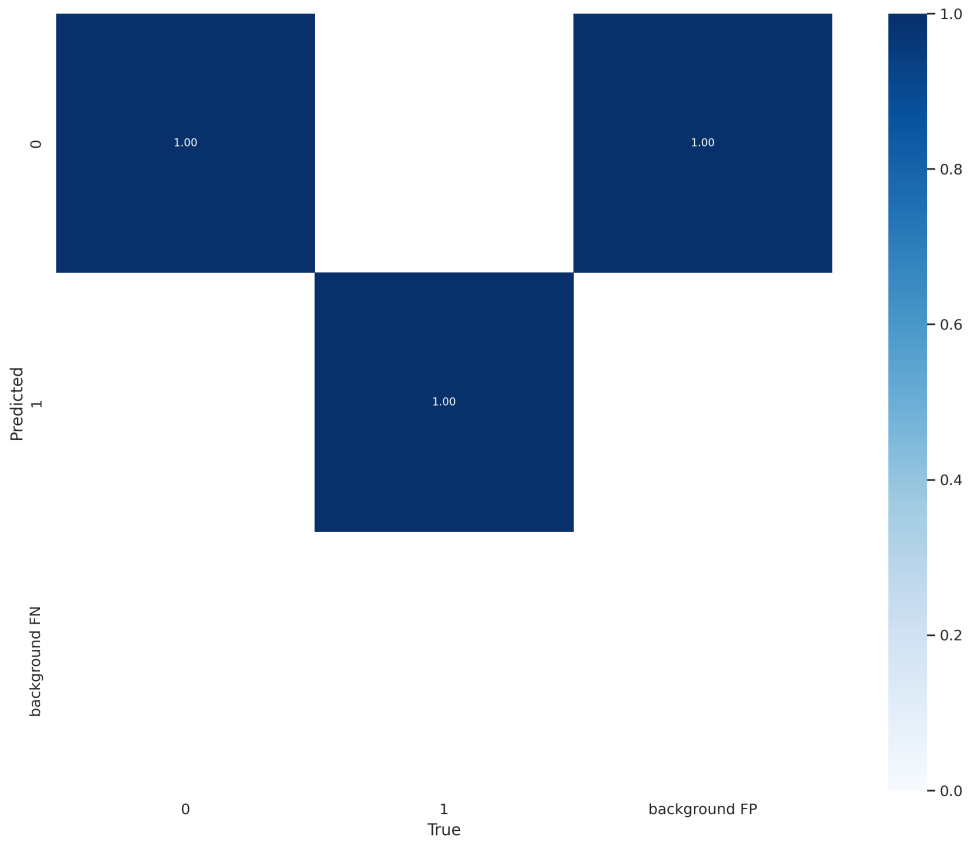

Figure 6.4: V3 Confusion Matrix with perfect recall and precision on validation set.

V3 is not overfitted since its validation classification loss does not increase after the initial convergence. It also attained a mAP score of 99.5% on the validation set, with perfect recall and precision (Figure 6.4) all within 25 epochs, though in hindsight it would have been preferred to train it for at least 45 epochs since a shorter training period may result in a more underfitted model. Regardless, the increased robustness of V3 (and the subsequent versions) is attributed to the training data augmentations and the training set expansion of additional general water expression cases.

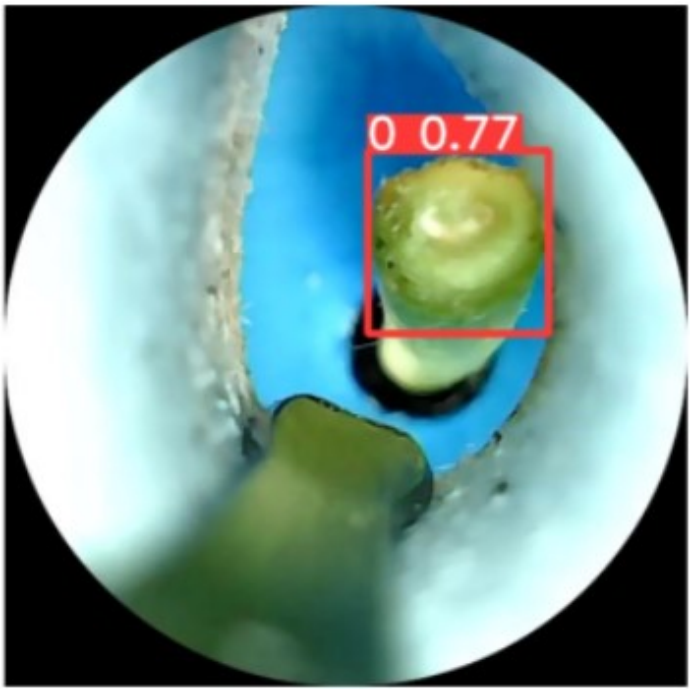

Figure 6.5: Correctly classified dry instance of the blurry case by V3.

V3 was given 119 test images from both the endocam and HQ Picamera configurations during simulated inference, and it was able to classify 118/119 test images correctly, a high success rate of 99.16%. The misclassified image is of the blurry case (Figure 6.5). Despite this, the model was still able to classify the other blurry case instances correctly, albeit with a lower confidence that ranges between $[0.68 - 0.80]$. In addition to classifying the majority of the blurry cases correctly, V3 was also able to classify the ambiguous case (Figure 6.1) correctly, the case that the previous versions struggled to recognize subtle changes between the dry and wet states. Another case seen frequently during our data collection that V3 (along with V4 and V5) was trained on is the bubbly case (Figure 6.6), where water expression resulted in a mass of white bubbles rather than a sap drop. This is a critical case to train our model on because the mass of white bubbles could potentially be

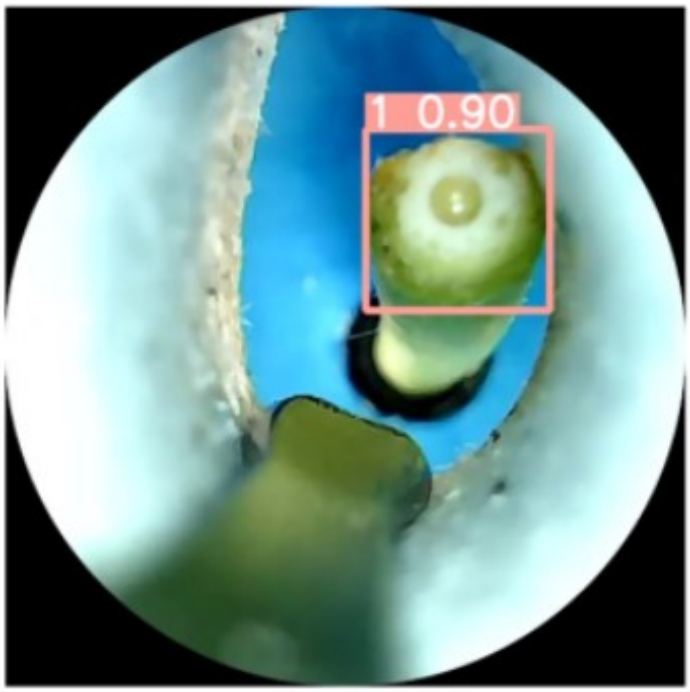

Figure 6.6: Correctly classified wet instance of the bubbly case by V3.

mistaken as the white pith of a dry xylem. However, misclassification on this case was not an issue for V3 since all instances of the bubbly case were classified correctly, with decent confidence around 0.9.

Of all test cases, the computer vision model performed the best on the HQ Picamera images (Figure 6.7), scoring $[0.91 - 0.95]$ confidence on most instances. This is due to the higher quality images the HQ Picamera is able to capture. Although all images from both camera configurations were resized to 640x640 prior to training, the downsizing of the high quality images retained better quality than the endocam images, so the model was able to classify the high quality images with more confidence.

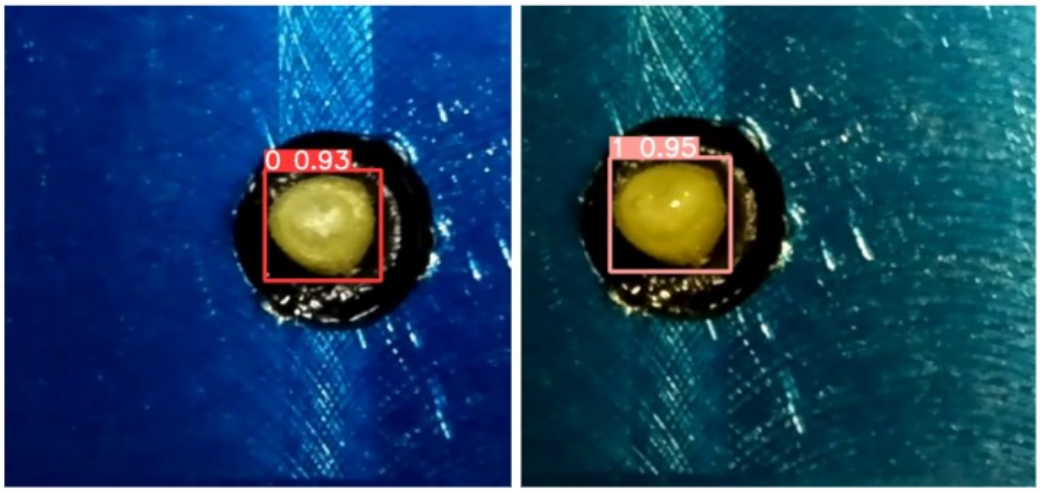

Figure 6.7: Correctly classified instances of the HQ Picamera images by V3.

6.2.4 Version 4

The fourth version, V4, is a slightly modified retrain of V3. The modifications include increasing the number of training epochs from 25 to 45 and applying Gaussian Blur as an additional pixel-level image augmentation. The intent of using the Gaussian Blur augmentation is to help the model recognize blurry or out-of-focus cases (better than V3) that may arise from faulty SWP measurement setup by a pressure chamber operator. However, V4 did not see the same performance (Figure 6.8) as V3 as the performance decreased slightly, with the success rate dropping to 114/119 (95.7%) images classified correctly. Additionally, the confidence for the correctly classified cases of V3 decreased substantially, with previously 0.9 confident dry cases dropping to only 0.6 confident. All misclassified images were of the blurry case, so the modification that created V4 seemed to have backfired. The following version, V5, omits the Gaussian Blur augmentation and opts for another, HSV augmentation.

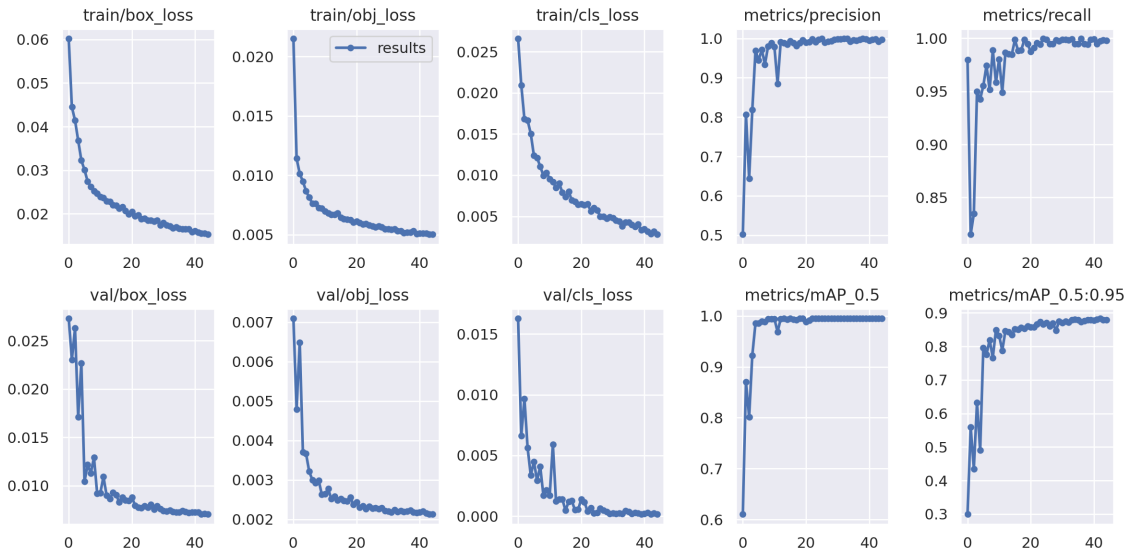

Figure 6.8: V4 Quality Metrics.

6.2.5 Version 5

V5 is also a slightly modified retrain of V3, with modifications including increasing the number of training epochs from 25 to 50, increasing the probability of applying specific augmentations, and applying the new HSV augmentation to the training data (Figure A.7). As mentioned earlier in Sec. 4.2, some notable features in the xylem through water expression include the appearance of a sap drop or bubbles, and color changes in the pith. The appearance of water is a color invariant feature, so HSV was applied to capitalize on this wet state sensitivity. As for pith color variation, HSV helps highlight dry state features more prominently, like the reddish-brown hues present in some xylems. The idea here is to use HSV to emphasize features characteristic to either state and have the model learn to recognize those distinctive features.

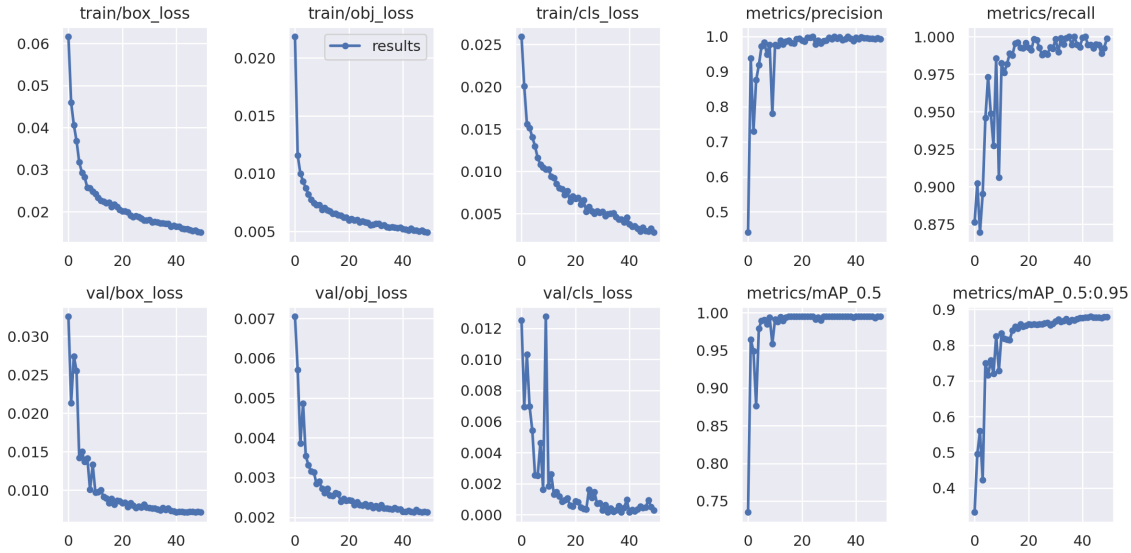

Figure 6.9: V5 Quality Metrics.

Initially, the resulting performance of V5 (Figure 6.9) over 50 epochs of training did not seem to yield results much better than that of V3. The classification confidence of all cases except for the blurry case remained the same, about 0.9 confidence in general. Contrary to our expectations, HSV augmentation initially did not seem to improve the performance of the model since the success rate was 114/119 (95.7%). Much like V4, all misclassified images were of the blurry case. However, a closer reexamination of the blurry case images revealed that the dry state of the blurry case does exhibit features characteristic to the non-bubbly wet state, such as light refraction and softer profiles. V5 classified the dry blurry case as wet with a confidence of approximately 0.9, and while this seems to be poor performance if considering metrics purely, the ability for V5 to recognize these features that are characteristically wet prove otherwise. Additionally, V5 scored high, but not ideal precision and recall (Figure 6.10), so it is less likely to be overfitted compared to V3.

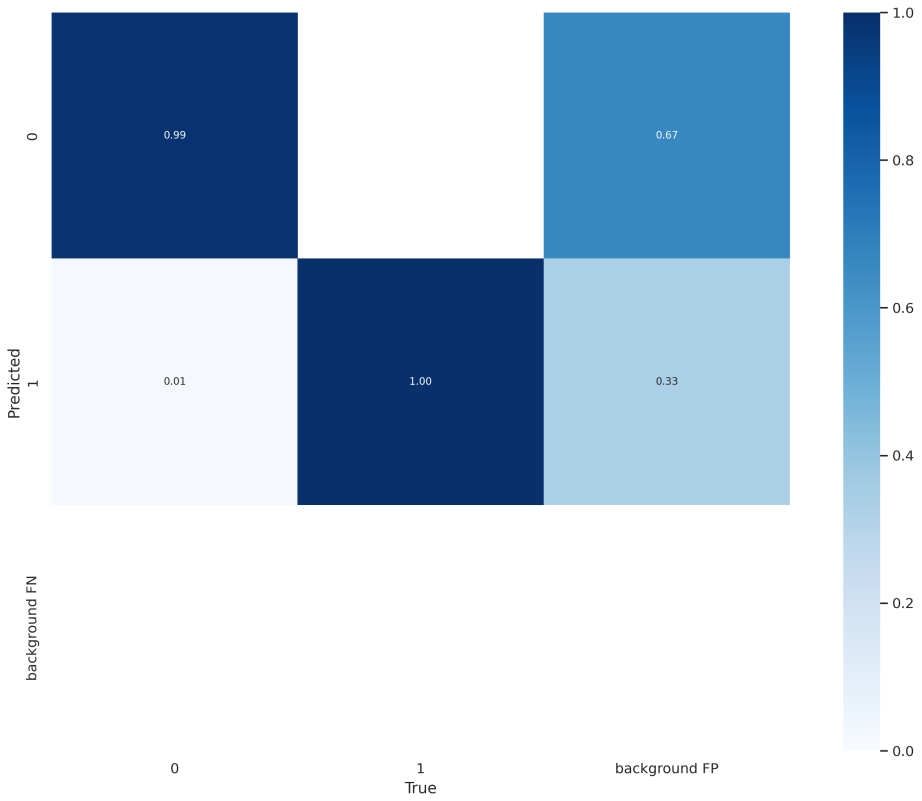

Figure 6.10: V5 Confusion Matrix.

6.3 Discussion

Of all versions of the computer vision model trained, V3 and V5 are the most promising given their classification performances. V3 had a success rate of 99.16% and V5 had a success rate of 95.7%. On paper it seems that V3 is the better model, but V5 is more consistent with identifying characteristically wet features (Sec. 6.2.5) and V3 identified the dry blurry cases correctly. Given this detail and the ideal precision and recall on the validation set, V3 may be slightly overfitted to recognize specific cases, and can't predict as well as V5 in the general sense. Regardless, both V3 and V5 will be used for our next evaluation steps including live inference with a static pressure chamber.

Chapter 7

Conclusions and Future Work

This chapter concludes the research presented herein and discusses some directions and ideas for future continuation of the automated SWP analysis project. Some future plans already discussed with the Agricultural Robotics (Agrobotics) project group include continued development and refinement of the system presented herein on a larger, static pressure chamber, and the integration of the system with the other projects on the Clearpath Robotics UGV platform, Husky[1].

7.1 Conclusions

We were successful in developing a computer vision model for a control system in the grander goal of automating the SWP analysis process. The final versions of the computer vision model, namely V3 and V5, attained mAP scores of 99.5% on the validation set and high precision and recall. V3 classified the testing set with a success rate of 99.16% and V5

[1]Husky UGV from Clearpath Robotics: https://clearpathrobotics.com/husky-unmanned-ground-vehicle-robot/

with a success rate of 95.7%. Both versions of the model have merit, with V5 possessing an edge over V3 due to its additional HSV augmentations, despite having a lower classification success rate. Between the two model versions, V5 also is more consistent in recognizing characteristically wet features in the dataset than V3. Our models are robust since they were capable of discerning dry and wet states from multiple different cases likely to be seen in a real-world scenario. This proves the validity of our system and its potential to automate the SWP analysis process.

7.2 Future Work

7.2.1 Inference and Development on Static Pressure Chamber System

The immediate future will involve running live inference with the models on a static pressure chamber setup and refining the system so detection can occur real-time. We emphasize the use of a static pressure chamber here because it will provide much needed stability when conducting SWP measurements, unlike the pump-up chamber used for data collection. Additionally, the static design offers higher maximum operating pressure limits to permit the analysis of a wider range of crops.

The general direction for development on the static pressure chamber system is to design a new camera mount for the HQ Picamera setup, and devise a method to feed pressure measurements to the Raspberry Pi. These pressure measurements can be collected from an embedded pressure sensor in the chamber, and they function as either the detected SWP value at water expression or a threshold to release air pressure from the chamber

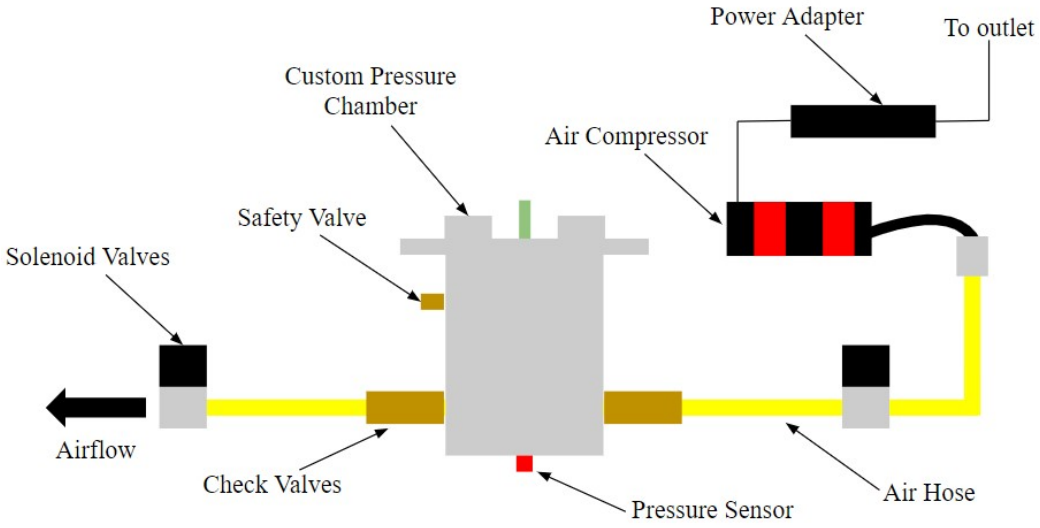

Figure 7.1: Static pressure chamber concept for computer vision control system to be integrated with.

to prevent overpressurization. The system concept is summarized in Figure 7.1, where the solenoid valves are controlled to permit airflow in or out of the chamber based on the readings from the pressure sensor and the detection state of the computer vision system.

7.2.2 Incorporation with other Agrobotics Systems

Once the live inference system functions and is refined, the computer vision control system will be incorporated with the other subsystems under Agrobotics, including the leaf cutter end-effector system through ROS[2] nodes. The intent is to have these systems coordinate with each other on the Husky UGV platform to enable in-situ field testing of a truly autonomous SWP measurement system.

[2]ROS (Robot Operating System): `https://www.ros.org/`

Appendix A

Supplementary Material

A.1 Version 3

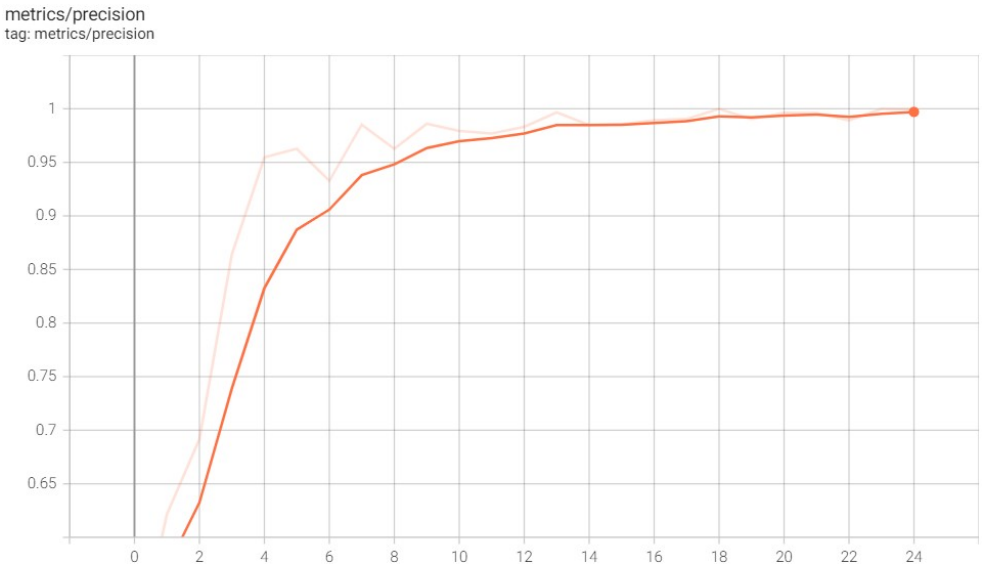

Figure A.1: V3 Precision.

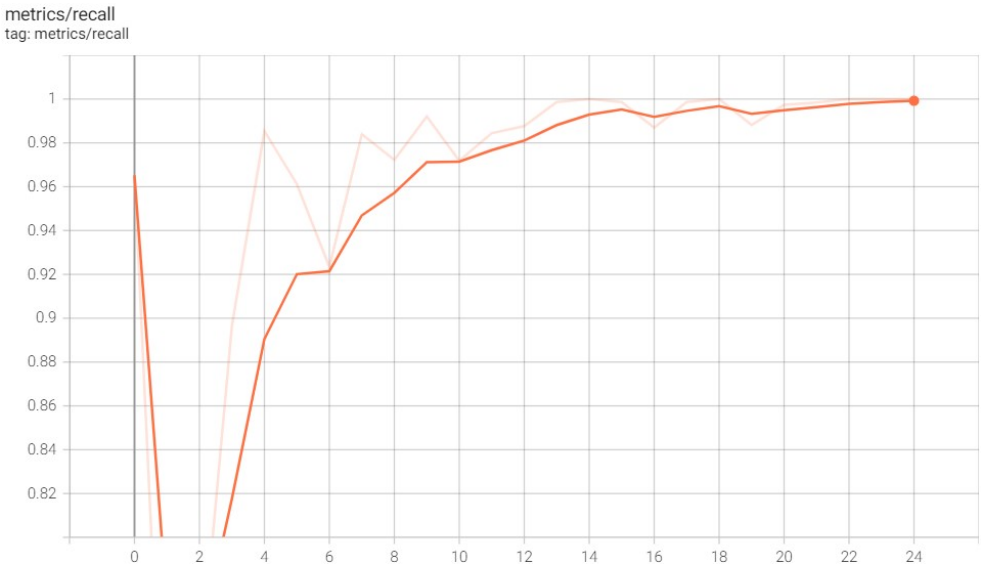

Figure A.2: V3 Recall.

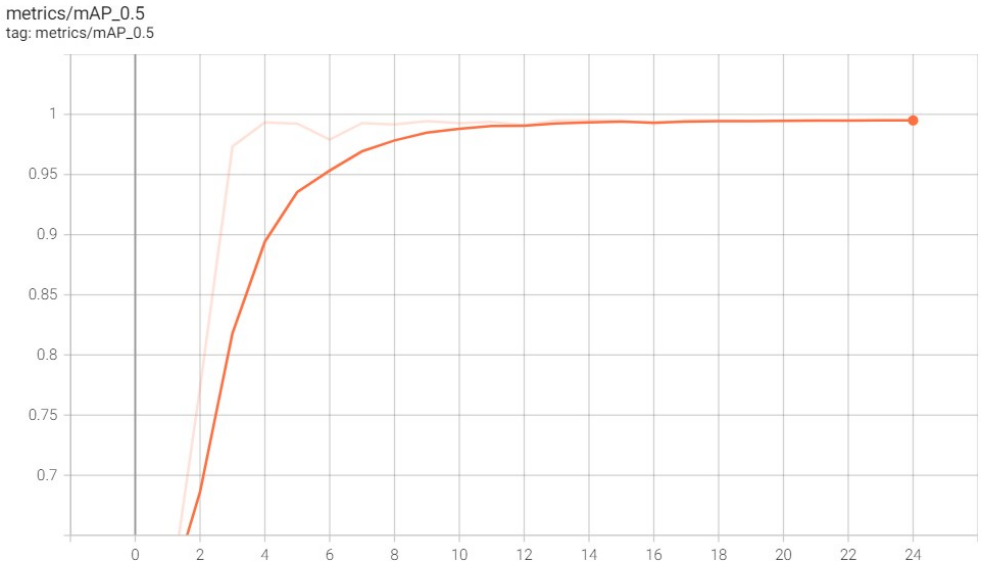

Figure A.3: V3 mAP Score.

A.2 Version 5

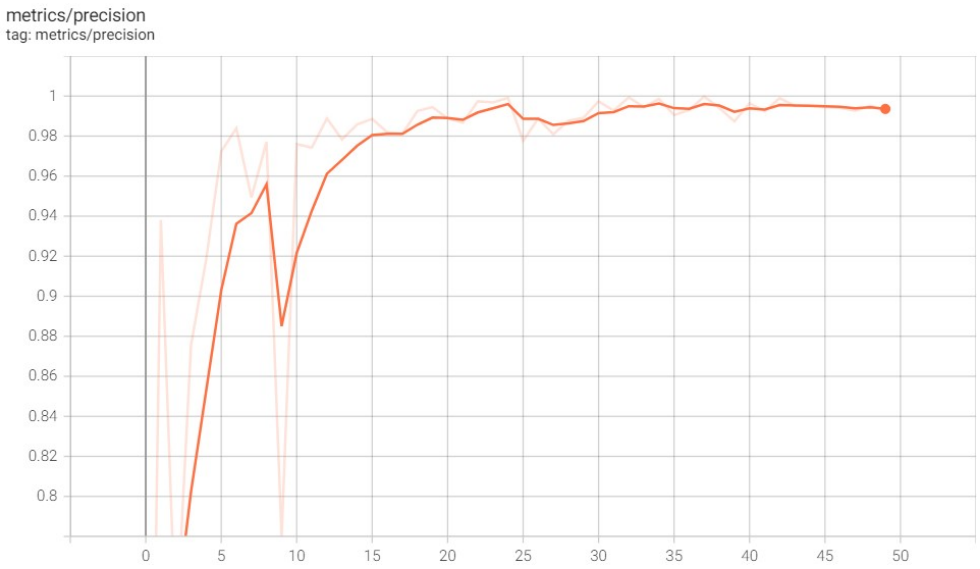

Figure A.4: V5 Precision.

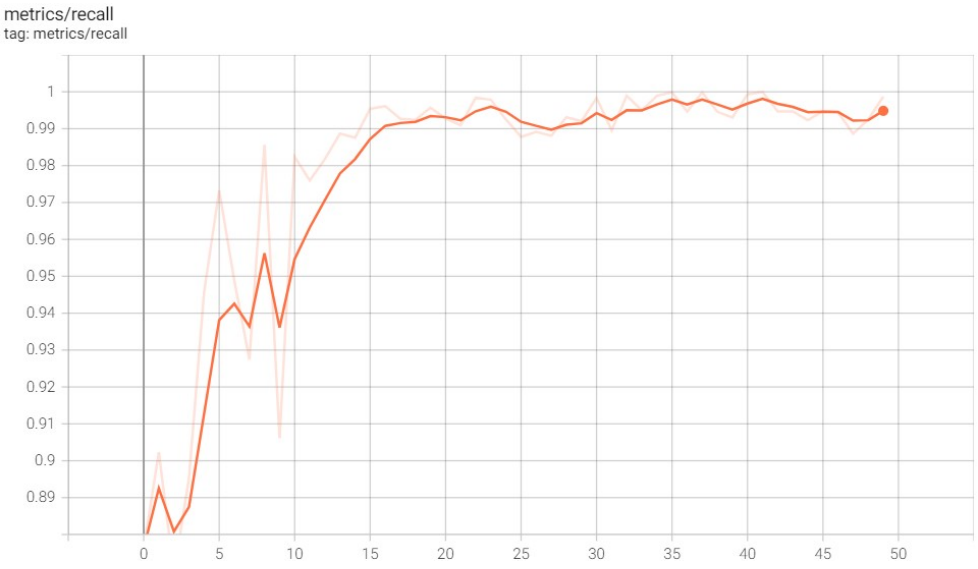

Figure A.5: V5 Recall.

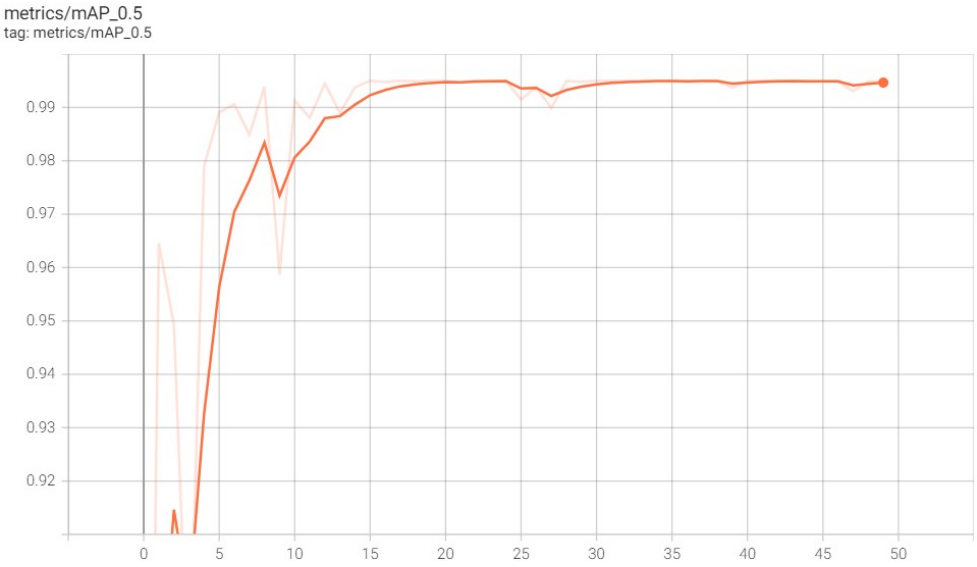

Figure A.6: V5 mAP Score.

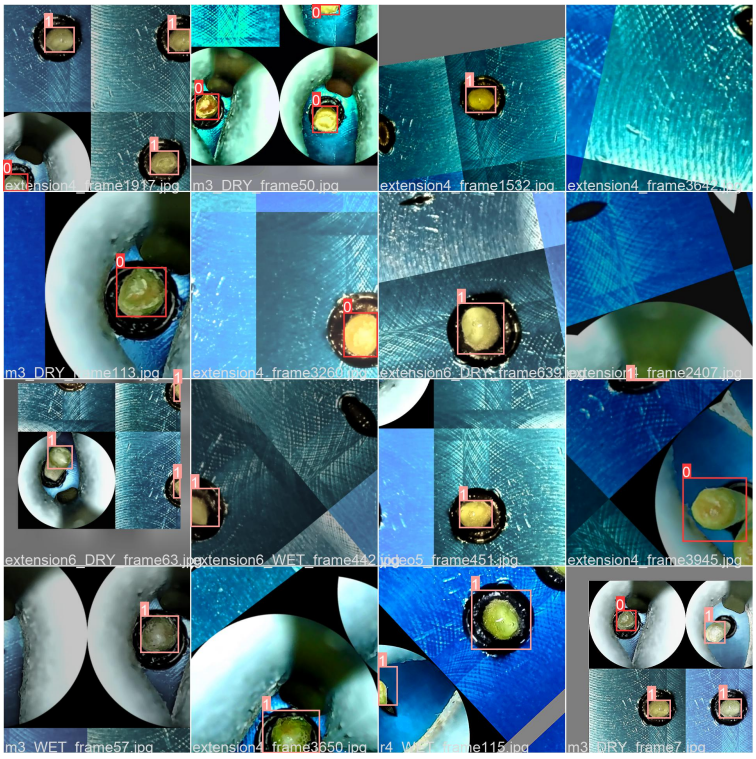

Figure A.7: Training batch images for V5.

Bibliography

[1] N. Zhang, M. Wang, and N. Wang. Precision agriculture—a worldwide overview. *Computers and Electronics in Agriculture*, 36(2):113–132, 2002.

[2] G. Vellidis, V. Liakos, W. Porter, M. Tucker, and X. Liang. A dynamic variable rate irrigation control system. In *International Conference on Precision Agriculture*, volume 31, pages 713–720, 2016.

[3] A. Fulton, J. Grant, R. Buchner, and J. Connell. Using the pressure chamber for irrigation management in walnut, almond and prune. *ANR*, 2014.

[4] A. Fulton, R. Buchner, C. Gilles, B. Olson, N. Bertagna, J. Walton, L. Schwankl, and K. Shackel. Rapid equilibration of leaf and stem water potential under field conditions in almonds, walnuts, and prunes. *HortTechnology*, 11(4):609–615, 2001.

[5] P. F. Scholander, E. D. Bradstreet, E. A. Hemmingsen, and H. T. Hammel. Sap pressure in vascular plants: negative hydrostatic pressure can be measured in plants. *Science*, 148(3668):339–346, 1965.

[6] M. T. Tyree and H. T. Hammel. The measurement of the turgor pressure and the water relations of plants by the pressure-bomb technique. *Journal of Experimental Botany*, 23(1):267–282, 1972.

[7] M. Meron, D. W. Grimes, C. J. Phene, and K. R. Davis. Pressure chamber procedures for leaf water potential measurements of cotton. *Irrigation science*, 8(3):215–222, 1987.

[8] N. C. Turner. Measurement of plant water status by the pressure chamber technique. *Irrigation science*, 9(4):289–308, 1988.

[9] J. M. Bennett. Problems associated with measuring plant water status. *HortScience*, 25(12):1551–1554, 1990.

[10] V. Saiz-Rubio, F. Rovira-Más, A. Cuenca-Cuenca, and F. Alves. Robotics-based vineyard water potential monitoring at high resolution. *Computers and Electronics in Agriculture*, 187:106311, 2021.

[11] P. Jain, W. Liu, S. Zhu, C. Chang, J. Melkonian, F. E. Rockwell, D. Pauli, Y. Sun, W. R. Zipfel, N. M. Holbrook, et al. A minimally disruptive method for measuring water potential in planta using hydrogel nanoreporters. *Proceedings of the National Academy of Sciences*, 118(23), 2021.

[12] P. F. Scholander, H. T. Hammel, E. A. Hemmingsen, and E. D. Bradstreet. Hydrostatic pressure and osmotic potential in leaves of mangroves and some other plants. *Proceedings of the National Academy of Sciences of the United States of America*, 52(1):119, 1964.

[13] J. S. Boyer. *Measuring the water status of plants and soils*. Academic Press, Inc., 1995.

[14] J. S. Boyer. Leaf water potentials measured with a pressure chamber. *Plant Physiology*, 42(1):133–137, 1967.

[15] G. A. Ritchie and T. M. Hinckley. The pressure chamber as an instrument for ecological research. In *Advances in ecological research*, volume 9, pages 165–254. Elsevier, 1975.

[16] J. S. Boyer. Relationship of water potential to growth of leaves. *Plant physiology*, 43(7):1056–1062, 1968.

[17] T. Zhao, B. Stark, Y. Chen, A. L. Ray, and D. Doll. Challenges in water stress quantification using small unmanned aerial system (suas): Lessons from a growing season of almond. *Journal of Intelligent & Robotic Systems*, 88(2):721–735, 2017.

[18] H. Vila, I. Hugalde, and M. D. Filippo. Estimation of leaf water potential by thermographic and spectral measurements in grapevine. *RIA: Revista de Investigaciones Agropecuarias*, 37:46–52, 2011.

[19] A. Patel, W. S. Lee, N. A. Peres, and C. W. Fraisse. Strawberry plant wetness detection using computer vision and deep learning. *Smart Agricultural Technology*, 1, 2021.

[20] A. Fuentes, D. H. Im, S. Yoon, and D. S. Park. Spectral analysis of cnn for tomato disease identification. In *Artificial Intelligence and Soft Computing*, pages 40–51. Springer International Publishing, 2017.

[21] Y. Sharon, B. Bravdo, and N. Bar. Aspects of the water economy of avocado trees (persea americana, cv. hass). *South African Avocado Growers' Association Yearbook*, 24:55–59, 2001.

[22] Y. Sudoh. Optical system for industrial camera that achieves both close minimum object distance and high resolution. *Proc. IS&T Int'l. Symp. on Electronic Imaging: Intelligent Robotics and Industrial Applications using Computer Vision*, pages 460 1–460 5, 2019.

[23] N. A. Switz, M. V. D'Ambrosio, and D. A. Fletcher. Low-cost mobile phone microscopy with a reversed mobile phone camera lens. *PLoS ONE*, 9(5), 2014.

[24] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao. Yolov4: Optimal speed and accuracy of object detection. *CoRR*, abs/2004.10934, 2020.

[25] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[26] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin. Albumentations: Fast and flexible image augmentations. *Information*, 11(2), 2020.

[27] H. Wang and Z. Song. Improved mosaic: Algorithms for more complex images. *Journal of Physics: Conference Series*, 1684(1):12–94, Nov 2020.

[28] G. Yadav, S. Maheshwari, and A. Agarwal. Contrast limited adaptive histogram equalization based enhancement for real time video system. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 2392–2397, 2014.