Image-based Microgreen Height-Yield Estimation

by

Jonathan Trevatt

School of Information Technology and Electrical Engineering,

University of Queensland.

Submitted for the degree of Bachelor of Engineering

in the division of Mechatronic Engineering (Extended Major)

NOVEMBER 2022

Faculty of Engineering, Architecture, and Information Technology

Jonathan Trevatt
42/15-17 Magellan Road
Springwood, QLD
4075

14/11/2022

Professor Michael Bruenig
Head of School
School of Mechanical and Mining Engineering
The University of Queensland
Queensland 4072

Dear Sir,

I, Jonathan Trevatt, am the author of the thesis entitled, Image-based Microgreen Height-Yield Estimation.

I submit this thesis for consideration as partial fulfilment of the Bachelor of Engineering (Hons) at The University of Queensland.

Upon The University of Queensland's acceptance of my thesis, the University and I agree that my thesis is to be kept confidential for a minimum period of five (5) years.

I warrant that no third-party copyrights or legal rights are infringed by my thesis.

All the work contained within this Thesis is my original work except where otherwise acknowledged.

Signature:

Jonathan Trevatt
Student ID: 43538706

Acknowledgements

I would like to thank my academic supervisor, Dr Mashhuda Glencross, and my industry supervisor, Andrew Hannam for their guidance in research direction, and for their help in producing a general outline for this report.

I would also like to thank my mother, Rachel Trevatt, for her efforts to help conduct and collect data for experiments, as well as my siblings, Daniel Trevatt and Katie Meiklejohn for their help with proof-reading and formatting.

Abstract

We attempted to develop a simple method of estimating the height and yield of radish microgreens using mobile phone cameras with the goal of increasing the accessibility of performing meaningful plant growth experiments for individuals and small business owners.

We produced a database of over 1000 images of radish microgreens, along with their corresponding height and yield measurements for use in this and future research.

We found a height-per-yield relationship over time of radish microgreens to be linear, with ((0.00011 * age in days) + 0.00324) grams per mm per shoot.

We used computer vision techniques and fiducial markers to estimate the height of microgreens, and from this result, an estimate of the yield.

We evaluated the algorithm on microgreens grown for commercial purposes, and it was found that it could not estimate the heights or yields within a minimum useful accuracy of ±20%.

Contents

List of Figures

1 Introduction

1.1 Problem to solve

It is challenging, time-consuming, and costly for small growers and individuals to take the consistent quantitative measurements that are required for monitoring and improving the efficiency of growing high-nutrient (healthy) foods, which consequently reduces the economic viability of the small grower.

1.2 Goals

1.2.1 Rationale for Goals

In recognition that the problem to solve is quite large and complex, I selected goals that, the pursuit of which, would produce valuable outputs for additional work. I approached the goals in consultation with the industry partner through their lens of creating practical next steps for further development.

1.2.2 Broad Goals

We posit that modern computing solutions can be used to assist in economic local production for small growers. To this end, the overarching goal of this project is to develop localised farming technology to benefit everyone nutritionally. The aims we work towards (but do not necessarily achieve in the scope of this research) are to develop general technology for plant health and yield data collection systems that (with further development) could:

1. Enable supervisory and **feedback control** for a small farming operation by improving data collection techniques available to them – increasing nutritional food security and profit by helping sustain a production steady-state.
2. Enable **economic experimentation** to optimise small farming operations without a high requirement for expertise – increasing food availability by increasing local production.
3. Have the potential to be expandable in capability.
4. Be **cheap and accessible** for a small farming operation.

To work towards achieving these goals, we propose a system using image processing techniques and computer vision to estimate yield.

1.2.3 Specific Goals

For the scope of this project, some more specific goals are:

1. Create a model that estimates yield (in grams) of microgreens with an error of at most 20% from images or video captured by a modern smart phone device. The industry partner commissioning this study estimates that this would be sufficient to have useful merit for their business as a first investigation*.
2. Evaluate and record extensions and limitations of the model and identify a path for improvement for future work.
3. Start a database of images and associated data that could be used for future study.

*This value comes from multiple business factors. 20% is a reasonable safety factor to fill a large order. At 20%, excess could be sold. For the purposes of detecting deviation from expected growth, a 20% difference is a minimum useful range.

2 Background

2.1 Rationale for estimate yields

There are many plant metrics that we could have chosen to investigate.

The vitamin and mineral content are particularly important for plants consumed as a healthy food. However, investigating vitamin and mineral content would require external lab tests, which are expensive and would not be within the budget for this thesis.

Plant colour is a strong indicator of health and can even be used to help diagnose the cause of underperformance. A study at Hankyong National University showed that a machine vision algorithm based on colour could distinguish between lettuce with poor or sound health with 80.8% accuracy [1]. A paper published by IEEE discusses the methods used to detect plant diseases using images of their leaves [2].

Early detection of pests or damage due to pests can allow farmers to act quickly to reduce damage to crops. A team in 2021 developed an automatic detection and counting system for pests that could correctly reflect trends for pest populations in large fields [3].

Yield is a desirable metric to study because:

- It is the primary metric for quantifying and valuing the microgreens for sharing or sale
- It can be used to help determine the best time for harvest [4]
- It can be used to indicate the general health of the plants, providing insights to help optimise growing conditions
- Addressing automatic yield estimation meets the thesis goal of enabling accessible experimentation on microgreens.

2.2 Rationale for methodology

2.2.1 Load cells vs cameras to measure yield

Load cells (the underlying sensor used by most scales) are used to measure yield after harvesting because they are cheap and available for many different maximum loads and precisions. But there are several reasons why load cells are unsuitable for measuring yield over time for many samples.

Load cells are prone to drift, partly due to environmental factors like temperature fluctuations. Drift can cause scales placed under a tray of microgreens to become inaccurate over time. Correcting drift can be achieved by using an identical load cell under a similar known weight in the same room. By comparison, camera images will not produce different results over time.

While individual load cells are inexpensive, having a load cell under each individual tray for even a small-scale farm would significantly increase costs.

A robotic system can be used to weigh trays of plants consecutively, but a similar system can be used to take images of them with a camera.

An advantage to using a camera over a load cell, is that while a load cell can only be used to measure forces, the data captured by a camera is much more versatile. That is, if you have an effective algorithm to interpret it.

2.2.2 Analytical vs artificial intelligence methods

For an analytical method, the algorithm is fully understood, because it is written by a human - but, it needs to be written by a human, which limits its complexity. Solutions developed using artificial intelligence (AI) methods (in this context, meaning neural networks and learning algorithms) are generally not as well understood, but can have the ability to take many inputs into consideration in building complex models. Another disadvantage of AI systems is that they often require vast quantities of tagged data to be trained. Particularly so for image analysis. For example, deep learning algorithms for recognising human faces require millions of images of faces tagged with features. The primary reason we chose an analytical method over AI is because we would have needed more datapoints of microgreens than I have access to or could produce in the allotted time. One of the goals of this thesis was to make collecting this sort of data easier. Therefore, we thought it made sense to start with a simple analytical model for this initial proof of concept. If it worked, the system could be used to collect data more easily for more complex systems, including AI-based systems. If it did not work, then this provides justification to try again using a more complex model.

2.3 CV methods for yield estimation

There has been much research in the field of automated agriculture. In recent years, there has been interest in computer vision-based measurement techniques. However, most of this research has focussed on large-scale agriculture – not small privately owned farms. Most modern techniques use tools and equipment that are either not financially accessible or not feasible for many small farming operations.

2.3.1 Examples

The 'Robotanist' [5] is a mobile platform for high-throughput crop phenotyping using stereo cameras and convolutional neural networks. Multiple impressive studies utilising state of the art techniques using this platform have been released. But the platform is unsuitable for a backyard or hydroponic farm due to size and cost.

Farmbot is a platform for automated backyard farming which utilises computer vision techniques to control weeds. But the platform lacks any active yield estimation capability.

2.3.2 What is commonly measured

Many inputs to a farming system can be cheaply monitored and/or directly or indirectly controlled with consumer-grade sensors and actuators.
For example:

Soil/Growth medium	Temperature (+/- 1°C [6])	CO_2/O_2 (aeration)	Compaction	pH [7]	Moisture [8]	Nutrient (various)		Particle size

Air	Temperature	CO2/O2	Pressure		Humidity [9] (±3% RH / ±0.3 ℃)		Flow rate	
Water	Temperature	CO2/O2 (aeration)	Pressure [10]	pH		Nutrient (various)	Flow rate	Droplet size

Figure 1 – Examples of parameters that can be monitored or controlled in a farming system

The intensity and frequency distribution of ambient and direct light [8] can also be measured and controlled.

For a large commercial farms, plant health, growth, and yield-can be assessed broadly using light-aircraft-based aerial photography in visible and infrared light [9] [10]. The nutritional value of plants can be estimated using spectrography and chemical analysis. This information can estimate and optimise yield (and, by extension, food production and profits) by adjusting inputs to the system. Total yields are often high enough that some portion of a crop can be dedicated to experimentation, and small losses are inconsequential.

Crops dedicated to experimentation to improve future yields cut into profits because farming at a small scale means that individual plants are more significant than farming at a large scale.

Considering a leafy plant (where the leaf is the crop) in a small farm, any empirical measurements of health, growth, or yield metrics of individual plants usually must be manually measured, which is time-consuming and destructive (to the plant). Because of this, there often no empirical measurements of any health or growth metrics throughout the crops growth cycle, and the yield (as weight) is measured at the time of harvest.

For example, ‘Natural Yield Pty Ltd’ does not record empirical data throughout the growth cycle of their crop. They weigh the net product to determine yield after harvesting [4].

There is a gap in the market for systems to monitor plant growth for small farms, and filling this gap has the potential to shorten the latency of control efforts from the length of an entire growth cycle.

Diseases, pests, nutrient and chemical composition, colour, and leaf dimension and count are measured using various sensing devices.

2.3.3 Sensing Devices

E. Tunca used UAV-based high resolution thermal and visible imagery to estimate bell pepper evapotranspiration [10]. Sébastien Cuq used infrared spectroscopy and chemometrics to assess macro and micronutrient concentrations in vine leaves and grape berries.

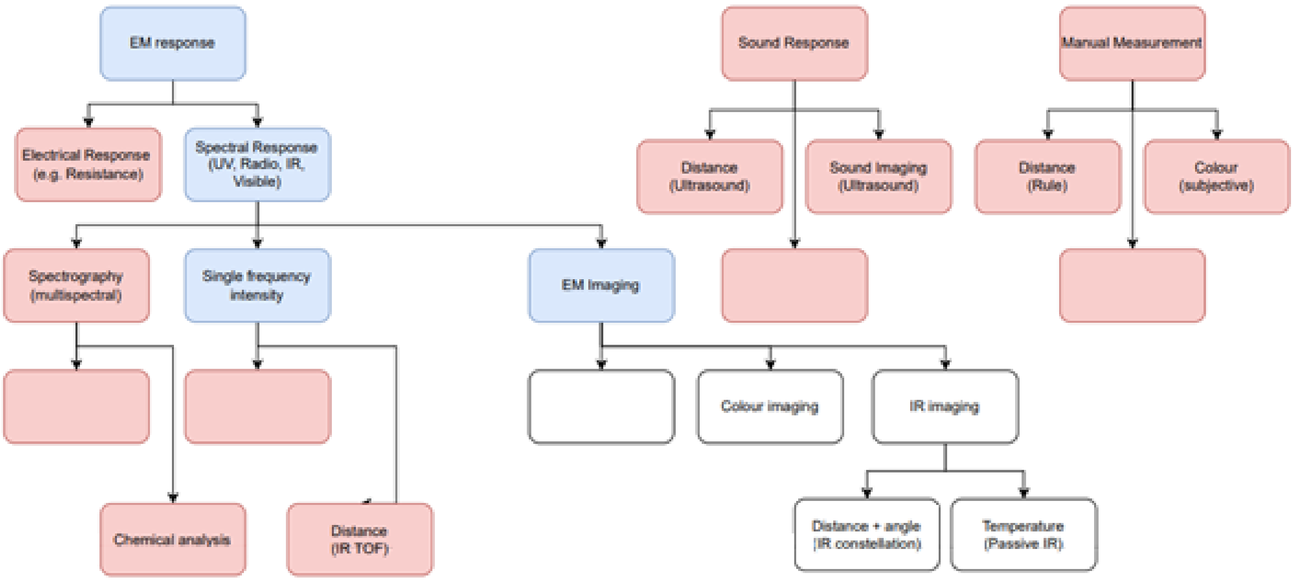

Figure 2 – Examples of different methods of taking measurements of plants

3 General Methodology

3.1 Assumptions being tested

The following are known assumptions. However, we may have made other assumptions without recognising them as such.

- We can use ArUco tags to correct perspective in mobile phone camera images.
- We can measure average microgreen height (in mm) using pixel measurements (from the perspective corrected images).
- We can correlate an average measured microgreen height with the current yield (in grams) of a sample if we know the number of shoots in the sample and the height-weight density of the plant.
- There are correlations between all known factors that we can use for estimation
- We can estimate the future yield of the microgreens by comparing their current height, weight, and age with historic records of the same.

3.2 Approach

I planned three main phases to the approach. The phases are not strictly linear but overlap. Each phase is described below.

Plan			Step		
			1	2	3
Phase	Analytic	1	Collect (classical) Data	Quantify Yield/height	Test Yield/height
	CV, CNN	2	Collect (Image) Data	Yield/height from Image (Manually)	Yield/height from Image (Automatically)
	Extend	3	Future yield from current (ideal case)	Future yield from current (non-ideal case)	Test limits

Figure 3 – Diagram describing experiment methodology phases

The first step to achieve the goal of estimating yield from height is to characterise the relationship between the yield and the height of the microgreens (in particular, for radish microgreens).

A potential challenge is that as leaves grow primarily in the horizontal direction, leaf growth estimation should consider this ideally. Thus, there is an inherent assumption to be tested that the height and leaf size correlate strongly, or at least that the average leaf size between pots is consistent enough not to significantly affect the height-yield relationship's precision.

Once a height-yield relationship is established and evaluated using manual height measurements, the next step is to assess the accuracy of estimating the height of the microgreens using pixel measurements from images. The pixel measurements can then be transformed to a yield estimation.

One of the thesis goals is to develop a system that laypeople can use. Thus, I used an algorithm to automatically correct the perspective from imperfectly captured images and estimate the yield. Finally, we want to explore the system's limits and evaluate whether the system can record growth data over time and accurately predict future yield and whether it can be used to detect sub-optimal growth patterns.

4 Experiments

4.1 Standard method for preparing seeds and harvesting

The trays used have dimensions of approximately 520 mm by 260 mm at the bottom, and 525 mm by 265 mm at the top, with a depth of 70mm. Trays are sanitised with a hydrogen peroxide solution before use. Each tray has an inner and an outer lining. The inner lining has a mesh base to allow roots to grow through, and the bottom lining is ribbed to allow water to flow under the inner lining.
Each tray is filled with approximately 1 kg of coconut coir. Each outer lining is filled with 600 ml of water. The coir should darken slightly in colour as it soaks up the water from below. Note that the soil should be damp but not wet, and a different water quantity may be required depending on the humidity. The coir is then patted and spread evenly to flatten the surface.

For each tray, 55 grams of radish seeds are measured out and spread evenly on the surface of the coir. Loose coir that has fallen into the bottom lining is then sprinkled to slightly cover the seeds. The surface is then sprayed with water until damp (approximately 100 ml per tray).
The coir surface is then weighted (5-10 kg) and left for three days to germinate. After three days, the new shoots should have lifted the weighted tray. At this point, the new shoots will be yellow. They will turn green within a day when exposed to light. The trays are then moved under grow lamps, watered twice a day, and under light for 16 hours a day.

Sprouts are harvested at any point from now, but the ideal time to harvest Microgreens is approximately seven days after planting, just before the 'true' leaves appear. Microgreens are harvested manually with a knife at about 1-2 cm above the level of the soil or coir.

Figure 4 – Tray, standard preparation: Coconut coir and seeds

Figure 5 – Tray, standard preparation: Seeds germinating under weight

4.2 Experiment 1

4.2.1 Goals

1. To gain experience (for the sake of future experiments) in developing and testing a method to investigate the relationship between height and yield of radish microgreens over time.
2. To collect images of microgreens with known heights, yields, and age to contribute to a database for computer vision analysis - mainly to measure the height of the microgreens from the images for yield estimation.

4.2.2 General Method

We divided the microgreens into twenty-two pots (samples) in a self-watering system under grow lamps.

We measured each microgreen sample's average height and exact weight (after draining water) daily.

We took images and videos of each sample from all angles. The light intensity at the base of each pot was also measured. We harvested the microgreens at the end of the experiment and recorded their yield.

We conducted the experiment and the data collection to provide experience in the challenges we would need to overcome. The yield and height measurements enabled calculating a yield-height relationship over time. The photos and videos of samples combined with this data were to contribute to a library for image analysis.

4.2.3 Method of Analysis

Since the weight measured for this experiment was of the entire pot (including roots and soil), we needed to find an estimate of the yield to calculate a relationship between height and yield.

To estimate the projected yield, we weighed an identical pot filled with soil, soaked it in water, and drained it. We took the measured weight of each sample. What remained was an estimate of the weight of the plants over time, including roots.

We assumed that the proportion of the yield versus the total weight of the plant is roughly constant over time and that this proportion is equal to the true yield (measured after harvesting) over the estimated weight of the whole plant immediately before harvesting. An estimate of the yield over time could then be calculated for each sample for each day and fit with a linear regression.

4.2.4 Procedure

The following steps were carried out to conduct the experiment.

4.2.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. An aluminium frame was built to hold two hydroponic channels, two grow lamps, and have space below for a reservoir.
2. Eleven evenly spaced circular holes (80 mm in diameter) were drilled in the lid of each channel to hold (86mm diameter by 75mm height) pots.
3. Each pot was filled to the same level with perlite stones as a growth medium (Figure 8).
4. The channels and reservoir were connected in a loop from reservoir to channel one, to channel two, then back to the reservoir.
5. Water was cycled through the reservoir, tubes, and channels using an aquarium pump.

6. A grow lamp was hung at the same height directly above each channel (Figure 9).
7. The lamps were set to automatically turn on every morning at 6:00 am and off at 10:00 pm each night.
8. Two trays of microgreens were prepared using the usual method (as above).
9. Once the weights were removed, each tray was divided into 18 equal circular samples using the cross-section cut from one of the cylindrical hydroponic pots (Figure 6).
10. The roots holding them to the rectangular tray inner lining were cut, and each sample 'puck' was placed on top of the perlite in a cylindrical hydroponic pot in the test rig (Figure 6, Figure 7, Figure 9Figure 7Figure 7).
11. The samples were labelled A1-A11 and B1-B11 for the 2 channels.
12. The light level (in LUX and PPFD) was measured at the base of each sample along the trays.

Figure 6 –Experiment 1: Top view of circular samples

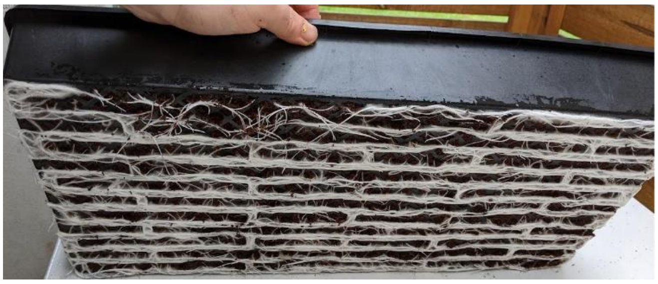

Figure 7 – Experiment 1: Microgreen roots growing through base of tray

Figure 8 – Experiment 1: Frame holding pots filled with perlite to equivalent levels

Figure 9 – Experiment 1: Frame showing grow lamps hanging above each plant row

4.2.4.2 Daily Procedure:

The following steps were completed for each day of the experiment for a total of 9 days, at approximately the same time each day.

1. Sample pots were lifted out and placed on a diagonal in their channel holes to drain.
2. Once they stopped dripping, sample pots were weighed individually on a set of kitchen scales and replaced in their holes.
3. The approximate average height of the shoots in each pot was measured (in 5 mm increments).
4. Each pot was photographed from above, behind, in front, and from both sides at various angles.
5. Each pot was videoed from angles all around.

4.2.4.3 Concluding Procedure:

The following steps were taken to complete the experiment.

1. At the end of the 9-day growth period, the microgreens were harvested and the actual yield for each sample recorded (in grams).
2. An 'empty' pot was prepared in the same manner as the other samples (with coir and perlite), except without any shoots, and was weighed.

4.2.5 Results

The data showed the height increase over time, as expected.

Once the outlier is accounted for, the difference between the minimum and maximum average heights ranged between 20 mm and 35 mm, and the average standard deviation is 7.4 mm.

The height growth rate slowed as the plants reached maturity and prepared to grow true leaves (Figure 11).

It was observed that just prior to the true leaves sprouting, the cotyledon stems would separate from each other. The true leaves would then sprout from the join between the cotyledons.

The light intensity at the base of each sample was between 1340 and 8610 Lux (Figure 15).

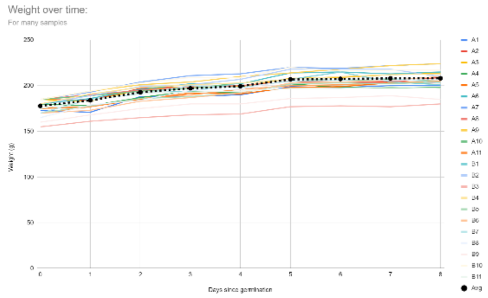

Figure 10 – Experiment 1: Weight over time.

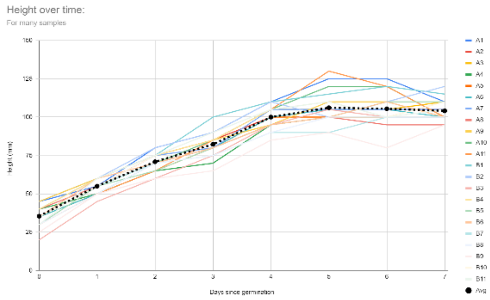

Figure 11 – Experiment 1: Height over time.

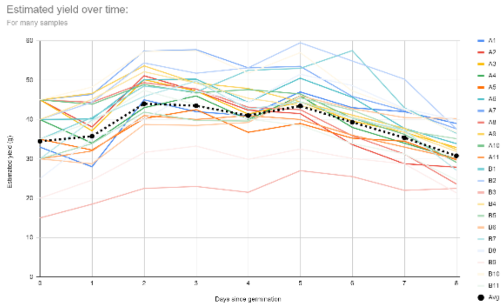

Figure 12 – Experiment 1: Estimated yield over time.

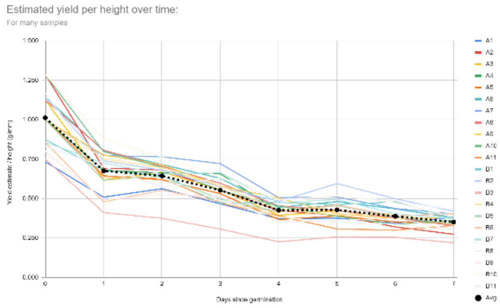

Figure 13 – Experiment 1: Estimated yield per height over time, yield-height correlation coefficient: 0.29.

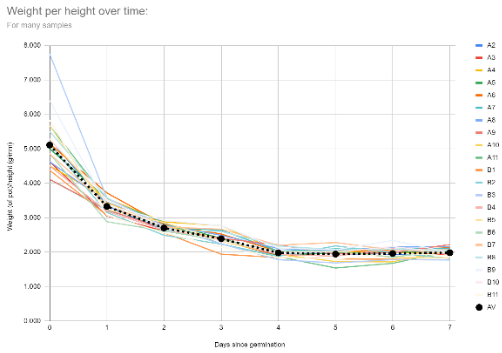

Figure 14 – Experiment 1: Weight per height over time, weight-height correlation coefficient: 0.77.

4.2.6 Analysis and Conclusions

We expected the rate of change of yield (g) and height (g) may decrease as the plant prepares to grow true leaves (towards the end of the experiment). The observation of the separation of the cotyledon stems just prior to the growth of the true leaves revealed two useful facts.

1. First, the observed decreased or even negative growth rate is due (at least in part) to an increased angle between the cotyledon stems. The increased angle separates and lowers the cotyledon leaves themselves and is caused by the true leaves sprouting between them. Our previous assumption was that the primary cause of the observed change was that the plants metabolic energy was being diverted from height growth to develop the true leaves.
2. Since the apparent change in height growth precedes the true leaves sprouting by about a day, it can be used as an indication that the true leaves are beginning to sprout. This is useful because true leaf growth on the microgreens is considered undesirable by the company sponsoring this research.

We expected that light intensity will correlate positively with growth rate. We observed the varying light intensity had a minor effect on the growth due to nature's ability to self-correct. The microgreens on the ends of the channels with lower light intensity, produced a slightly lower yield, however they tended to be taller and leaned slightly sideways towards the stronger light.

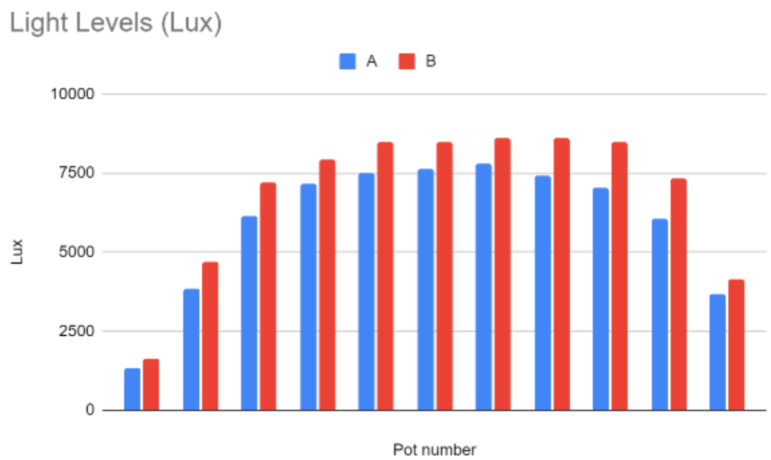

Figure 15 – Light intensity (in lux) as measured at each pot (sample) in groups A and B. Samples are listed relative to each other in the same positions as they were physically taken. I.e., the left-most sample in A was measured at the position of the left-most pot under the light over sample group A.

4.2.6.1 Assumptions

Some assumptions made in the development of this experiment:

- The difference in growth rate between samples due to difference in height over time is negligible compared to the difference due to the position of the samples along the length of the grow lamps (greatest intensity near the middle of the lamp).
- The difference in growth (as compared to real-world conditions) due to damaging the root's during transplantation is negligible.
- The proportion of root mass in a sample compared to harvestable mass of the shoots is stable over time and can be measured after harvest.
- The plants being watered constantly with flowing water will not significantly affect the growth (as compared to the periodic watering in commercial growth conditions).

4.2.6.2 Challenges and observations:

- Yield estimation accuracy is low because the sample's weight due to the yield is small compared to the weight of the sample due to the pot, perlite, and retained water. And because it was impossible to ensure each sample retained the same amount of water.
- Despite best efforts to control the light intensity variable, the light intensity experienced at each sample varied widely. This was taken as an opportunity to study the effect of differing light intensity on the microgreen growth.
- Creating the sample 'puck' required cutting the roots and transplanting sections into the pots. The effect of cutting the roots (if any) is unknown. This method was chosen over sowing the seeds directly in the pots so that weight could be applied to the seeds for the first 3 days, keeping the growth more consistent with the commercial growing conditions.

4.3 Experiment 2

4.3.1 Goals

1. To investigate the relationship between height and yield of radish microgreens, while overcoming some of the challenges presented during

experiment one. Particularly the challenge of needing to estimate the yield using a measurement of the weight of all the pot’s contents.

2. To use growing conditions that are more like those in which the plants are grown commercially than those used in experiment 1.
3. Record the density of the coir, and of the radish seeds (volumetric and packing density). Knowing seed density should allow us to estimate the number of seeds in a weighed sample.
4. To collect images of microgreens with known heights, yields, and age to contribute to a database for computer vision analysis. Particularly, to measure the height of the microgreens from the images for the purpose of yield estimation.

4.3.2 General Method

Instead of using pots in a hydroponic system to grow samples, the same trays as are used commercially were used, and the plants were maintained alongside their commercial counterparts. This was to make the growing conditions as similar as possible to the commercially grown plants.

Twenty-four sample groups were prepared across four different trays.

Some sample groups were harvested each day.

Heights were measured each day of the experiment for each sample group.

Yield was measured directly using which ever sample groups were harvested each day, illuminating the need to make yield estimates. This method also enabled us to have more than one sample group in each pot, allowing them to be grown under conditions more comparable to those for the microgreens grown for sale by the company (in trays). A downside to this method was fewer samples for the same quantity of microgreens grown. Each sample produced a single datapoint for yield and height when harvested instead of one for each sample for each day.

4.3.3 Method of Analysis

The height, yield, and yield per height over time was plotted. Because there was a large number of samples, a candlestick plot was used to summarise the data, with the bottom wick representing the minimum value in the population, the top wick representing the maximum value in the population, and the candlestick itself representing plus and minus one standard deviation from the norm.

4.3.4 Procedure

The following steps were carried out to conduct the experiment.

4.3.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. Four trays were prepared using the standard method.
2. The trays were each divided into 6 equal sections, creating 24 total samples.

4.3.4.2 Daily Procedure:

The following steps were completed for each day of the experiment from day 4 after planting.

1. The weight of each tray was measured and recorded.
2. The minimum and maximum height of the microgreens in each of the 24 sections was measured and recorded, as well as a subjective estimate of the average.
3. Photos were taken parallel and centred to the tray and distanced to capture the entirety from above and from the front (short side) (Figure 16, Figure 17)
4. A photo was taken of all trays together - from the side, so that all the shoots can be seen (except those obscured by the aluminium frame).
5. Three sections (1/2 a tray) were harvested, and their yields were logged.
6. The yield for each harvested segment was spread over a sheet of white paper minimising overlap as much as possible, and a photo was taken from above (Figure 19).
7. Trays were watered (trying to keep the same amount each time and watered around the same time each day - amount of water added was recorded.

4.3.5 Results

The plots of the height (Figure 20,Figure 21), yield (Figure 22), and height per yield (Figure 23) over time are shown. The height increases over time until peaking at approximately 90 mm on day 8. After day 8, the average height started decreasing slightly each day (Figure 20). The harvested yield each day increased until day 6, when it peaked at approximately 70 grams per sample. The harvests then decrease to approximately 50 grams on day 9, and then back up to 75 grams on day 11 (Figure 22). The yield per height initially increases to approximately 0.85 g/mm on day 6, before decreasing to 0.6 g/mm on day 9, then up again to 0.7 g/mm on days 10 and 11 (Figure 23).

Examples of images taken for the image database:

Figure 16 – Experiment 2: Image of sample 2-2, Front

Figure 17 – Experiment 2: Image of sample 2-2, Above

Figure 18 – Experiment 2: Image of sample 2-2, Above Front

Figure 19 – Experiment 2 – Image of a harvested sample

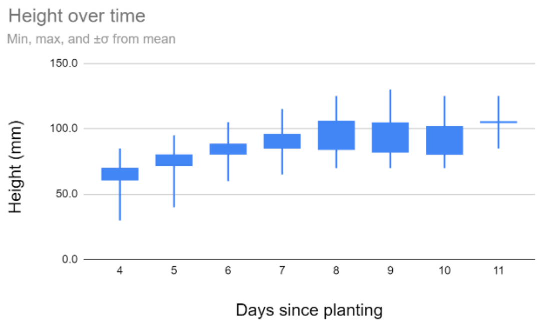

Figure 20 – Experiment 2: Candlestick summary of height over time. The bottom wick refers to the minimum value; top wick refers to the maximum value, and the candlestick refers to the +/- 1 standard deviation from the mean.

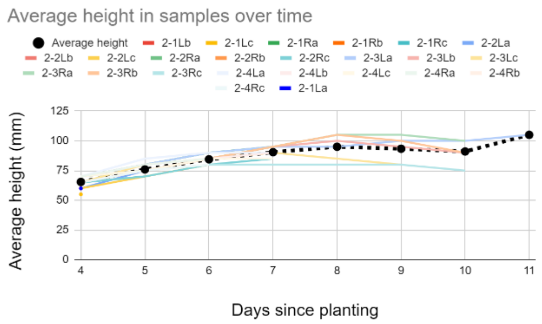

Figure 21 – Experiment 2: Line graph of average height in samples

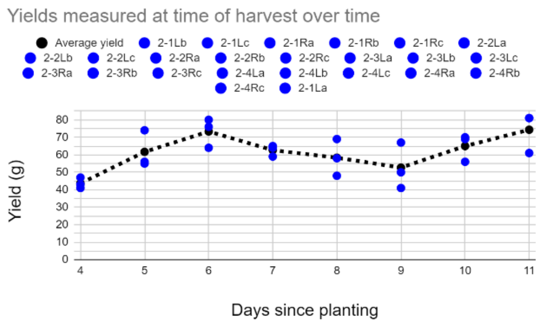

Figure 22 – Experiment 2: Yield measured at time of harvest over length of experiment

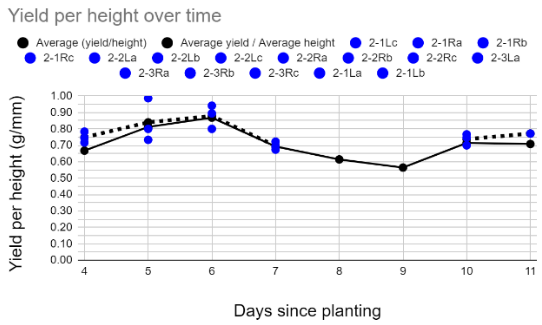

Figure 23 – Experiment 2: Yield per height of microgreen over length of experiment

4.3.6 Analysis and Conclusions

We initially kept the amount of water given consistent throughout this experiment, but there were a couple of days with higher temperatures and lower humidity, leading to the greens drooping from not enough water. After adding more water, they ‘perked up’, but some tangled as they continued to grow. Because the tangled shoots would have made it difficult to accurately measure height, these sections of shoots were harvested next. We believe that this sequency of events, along with not having enough samples to capture the variability of the yield may be the cause of the fluctuations in the yield measurements.

Measuring the minimum height of a sample proved to be a difficult and ultimately fruitless task. The shortest shoots were hidden under a canopy of taller shoots since some seeds never sprouted and some shoots would never grow to a harvestable size. We suggest that this metric is not worth recording in future experiments.

Photographing the harvested yield on a sheet of paper presented challenges. The task of separating shoots to prevent occlusion for the sake of future analysis proved arduous and was not completed. Shoots that are separated and placed on the paper are easily bumped or blown over each other.
For the sake of accurate measurements in future analysis, it was ideal for images to be taken with minimal perspective or lens distortion, and with even lighting. This is not a task that is quickly or easily accomplished without a dedicated setup for this purpose.
To solve some of these problems, we recommend a future experiment to use tape to secure the samples to paper, and a flatbed scanner to image them. This should eliminate lighting, lens, and perspective issues (though it would still be a time-consuming task).

Weighing the entire sample to estimate yield was inaccurate because the water retention in the soil was inconsistent. Weighing individual sprouts required removing them from the soil and repotting them, which damaged them and invalidated data for any subsequent growth.

There was some inconsistency in measuring the height because the soil (coir) was not perfectly flat.

We recommended that future experiments investigate the viability of using soil-free methods to grow the microgreens to allow individual shoots to be pulled out and weighed accurately before being returned without affecting their growth.

4.4 Experiment 3

4.4.1 Goals

1. To quickly assess the viability of using a soil-free plastic seed sprouting kit to measure the yield of individual shoots for later experiments.

2. To measure the density of radish seeds (i.e., the radish seeds from the brand and batch that we were using).

4.4.2 General Method

Three sample groups were prepared in soil-free sprouting kits and grown under identical conditions. One group was left unhandled. One group had samples (individual shoots) removed, measured, weighed, and returned each day. The final group had samples (individual shoots) removed, measured, weighed, and returned for the first two days, then left unhandled for the rest of the experiment.

This was to give rough indication of how severely the growth was affected by handling.

The weight-volume density of the coir and of the radish seeds were measured. The radish seed density was measured using the water displacement method. These were measured in case they became useful in future experiments.

4.4.3 Method of Analysis

This experiment was intended to be a quick test to assess the viability of the method. Since it was obvious early on by qualitative observation that the method would not be viable, no further analysis was required.

4.4.4 Procedure

The following steps were carried out to conduct the experiment.

4.4.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. A sprouting kit which consisted of stacked plastic trays 150 mm in diameter, was set up. Each tray had grooves in the bottom to hold water, and a hole set at a height to drain excess water.
2. Three sample groups (A, B, and C) were prepared with the same seed density as was used in previous experiments. Each tray represented a sample group (Figure 25,Figure 26,Figure 27).

3. The sample groups were placed side-by-side, receiving similar light to previous experiments.

4.4.4.2 Daily Procedure:

The following steps were completed for each day of the experiment.

1. The samples were watered twice daily.
2. Group A was left untouched.
3. Group C sprouts were individually removed, their heights and weights measured, then replaced in the tray.
4. Group B sprouts were individually removed, their heights and weights measured, then replaced in the tray for two days, then left untouched for the rest of the experiment.

4.4.5 Results

Coir density:	$\approx \frac{12kg}{50L} = 0.24\ kg/m$
Radish seed density (packed):	$\approx 0.9\ ml/g$
Radish seed packing ratio:	$\approx 69\%$

Figure 24 – Experiment 3: Measured density of coir and radish seeds

It was clear early in the experiment through qualitative observation that even the radish shoots in the group that was not handled (group A) did not grow similarly to those grown in soil. The shoots were falling over, and their growth was stunted.

Roots were tangled, but the shoots did not support themselves.
As the experiment went on, the more a sample group was handled, the worse it performed.

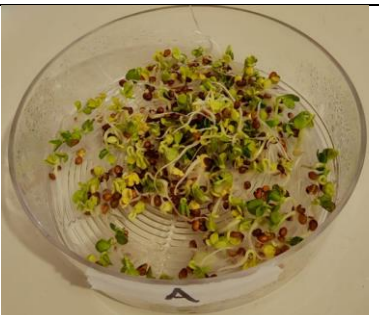

Figure 25 – Experiment 3: Sample A, day 3

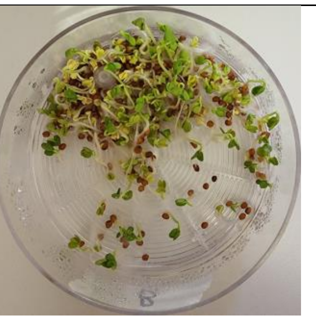

Figure 26 – Experiment 3: Sample B, day 3

Figure 27 – Experiment 3: Sample C, day 3

4.4.6 Analysis and Conclusions

The seed density was high enough that the roots tangled, which caused them to become damaged when separated to be measured.
Concurrently, the seed density was not high enough for the root tangles to form a mat that would support the shoots, which resulted in them falling over.
The sprouts did not grow as tall as expected, in part because they would fall over, and the stem would change direction to grow towards the light. Group B was unhealthy and small and some of the sprouts died. The entirety of group C died before the end of the experiment.

If the planting density were high enough for the shoots to support themselves and grow normally, then it would be impossible to separate the shoots for measurement without significantly damaging them and invalidating further measurements.

It was determined that this would not be a viable method to measure the yield of individual shoots over time, and that other methods would need to be explored instead.

4.5 Experiment 4

4.5.1 Goals

1. To assess whether yield and height measurements of radish microgreens grown using a steel mesh could be used to estimate the height-yield relationship for microgreens grown in coir.

4.5.2 General Method

A tray was prepared with a sheet of wire mesh instead of coir. Seeds were placed on the wire mesh in a grid pattern to enable consistent measurement of the height of individual shoots. Seeds required watering twice a day to prevent drying out.

4.5.3 Method of Analysis

This experiment was intended to be a quick test to assess the viability of the method. Since it was obvious early on by qualitative observation that the method would not be viable, no further analysis was required.

4.5.4 Procedure

The following steps were carried out to conduct the experiment.

4.5.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. A tray was prepared using a sheet of wire mesh instead of coir.
2. Water was poured into the lower tray.
3. The seeds were placed on the mesh in a grid pattern spaced at about 10mm between seeds using a template. The tray had 210 seeds divided into 6 groups of 35 samples each. Each sample represented a single seed (Figure 28).
4. A weighted tray was placed on top.
5. The seeds were sprayed twice daily with water to keep them moist, then the weighted tray was replaced.

6. The weighted tray was removed on day three and the seedling tray was placed under lights.

4.5.4.2 Planned Daily Procedure:

These steps were not carried out because the experiment was concluded early.

1. Once the roots reach the lower tray, water as normal.
2. Measure and record the height of each of the microgreen seedlings (from 5 days after planting).
3. Harvest at least one section of seedlings. (from 8-12 days after planting).
4. Weigh the total harvested part from each sprout.
5. Cut off and weigh the two cotyledons/leaves of each sprout.

4.5.5 Results

We observed that some of the seeds were damaged after weight was removed, and that many of the shoots grew crooked (Figure 29).

Figure 28 – Experiment 4: Seeds sprouting on steel mesh

Figure 29 – Experiment 4: Shoots on steel mesh with crooked stems, some fallen over

4.5.6 Analysis and Conclusions

We believe that the damage to the seeds was due to the weight added before germination because the steel mesh is a much harder surface than coir. When weights are added to germinating shoots growing in coir, the seeds are pressed into the soil and the pressure is evenly distributed on the seeds. When weights are added without soil, the force is concentrated on the contact points between the seeds and hard surfaces, crushing the seeds.

4.6 Experiment 5

4.6.1 Goal

The main goal of this experiment was to determine the validity of growing separate microgreen shoots in coir. The purposes of this are:

1. To measure weight of the microgreens more accurately at different ages and heights by directly measuring samples (rather than using interpolation to find estimates).
2. To measure weight of the microgreens more accurately at different ages and heights by using much more precise scales (+-0.001g) to measure individual sprouts rather than the average of groups of sprouts (thus collecting many more samples).
3. To collect images of microgreens with known heights, yields, and age to contribute to a database for computer vision analysis. Particularly, to measure the height of the microgreens from the images for the purpose of yield estimation.

4.6.2 General Method

We prepared a tray using the standard method, except seeds were placed in a spaced grid pattern to enable consistent measurement of the height of individual shoots. By collecting data for each individual shoot, we would be able to collect many more samples than in previous experiments.

More samples mean a better estimate of the yield per height, which works to achieve the experiment goals. We didn’t use this method in earlier experiments because it is highly time-consuming. The results of the previous experiments showed that it was necessary.

A potential drawback of this method is that the yield measured at different ages does not represent the same samples (since measuring the yield kills the plant). Also, the seed density is much sparser than in production (to make it easier to identify individual sprouts).

We assumed that the sample sizes were large enough that the average measurement of their weight at a given age would be representative of the population and that the decrease in seed density would not significantly affect growth.

4.6.3 Method of Analysis

This experiment was intended to be a quick test to assess the viability of the method. It became obvious early on by qualitative observation that using this method in conjunction with replacing the weight each day would not be viable, so no further analysis was required.

4.6.4 Procedure

The following steps were carried out to conduct the experiment.

4.6.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. A tray was prepared using the standard method.
2. The seeds were placed in a grid pattern spaced at about 10mm between seeds using a template. The tray had 210 seeds divided into 6 groups of 35 samples each (Figure 30).
3. Cardboard dividers were placed to help keep the sections separate however they had to be removed while the tray was under weight, and again later as they went soggy (Figure 30).

4.6.4.2 Daily Procedure:

The following steps were completed for each day of the experiment.

1. Photos were taken from day 3 (and put back under weight until day 5).
2. Height measurements were taken.

4.6.4.3 Planned Daily Procedure:

These steps were not carried out because the experiment was concluded early.

6. Once the roots reach the lower tray, water as normal.

7. Measure and record the height of each of the microgreen seedlings (from 5 days after planting).
8. Harvest at least one section of seedlings (from 8-12 days after planting).
9. Weigh the total harvested part from each sprout.
10. Cut off and weigh the two cotyledons/leaves of each sprout.

4.6.5 Results

We observed that many of the sprouts grew crooked which gave an inaccurate height measurement. The growth progression can be seen in Figures 21 to 24.

Figure 30 – Experiment 5: Seeds in tray, Day 1

Figure 31 – Experiment 5: Sprouts in tray, Day 3

Figure 32 – Experiment 5: Sprouts in tray, Day 4

Figure 33 – Experiment 5: Sprouts in tray, Day 5

Figure 34 – Experiment 5: Sprouts in tray, Day 6

4.6.6 Analysis and Conclusions

We are still uncertain whether the decreased seed density had a major effect on the growth. Despite being grown in soil as opposed to steel mesh, we believe the shoot damage was mainly due to the weight being removed and replaced each day for photos. The weight was kept until day five instead of day three, hoping it would cause the individual shoots to grow stronger.

However, the result was that the activity of replacing the weight damaged the shoots. We decided for future experiments only to use the weights for a maximum of 3 days.

3D printed dividers would be better – not too tall because of restricted airflow, and also can't be used while seeds are under a weighted tray.

-

4.7 Experiment 6

4.7.1 Goal

This experiment had multiple goals:

4. To measure weight of the microgreens more accurately at different ages and heights by directly measuring samples (rather than using interpolation to find estimates).
5. To measure weight of the microgreens more accurately at different ages and heights by using much more precise scales (+-0.001g) to measure individual sprouts rather than the average of groups of sprouts (thus collecting many more samples).
6. To prepare for the possibility of using more complex modelling of the sprouts if required, by developing a database of flatbed scanner images of the sprout stems and leaves along with their growth and weight records.
7. To collect images of microgreens with known heights, yields, and age to contribute to a database for computer vision analysis. Particularly, to measure the height of the microgreens from the images for the purpose of yield estimation.

A potential drawback of this method is that the yield measured at different ages do not represent the same samples (since the act of measuring the yield kills the plant). Also, the seed density is much sparser than in production (to make it easier to identify individual sprouts).

We made the assumptions that the sample sizes are large enough that the average measurement of their weight at a given age will be representative of the population, and that the decrease in seed density will not significantly affect growth.

4.7.2 General Method

We prepared trays so that we could measure the height of individual shoots over time. The yield of each shoot (the total and the leaves individually) was measured when harvested.

By collecting data for each shoot, we collected many more samples than in previous experiments. More samples mean a better estimate of the yield per height, which works to achieve the experiment goals. We didn't use this method in experiments before Experiment 5 because it is highly time-consuming. The results of the previous experiments showed that it was necessary.

We measured the thickness of the stems and leaves. Each sample was attached to a page, labelled, and scanned with a flatbed scanner. These measurements enabled more complex modelling for further research to improve the works of this thesis. This data also works towards the goal of developing a useful database with images of microgreens and measurements. Other use cases include image recognition modelling and phenotyping research.

4.7.3 Method of Analysis

We collected the height, weight (total and weight of just the leaves), and thickness of the leaves and stems (where measured) for each sample in a table.

The data was filtered to remove samples that did not grow to a harvestable height (defined in this experiment as any samples that did not exceed 5mm or had no leaves). This left 385 out of 420 samples (i.e., 92% of the seeds were viable).

The contribution of the yield due to leaf matter and stems were calculated separately.

We calculated the Pearson product-moment correlation coefficient for the height-yield relationship and the relationship between the yield contribution from stems and leaves.

We recorded the mean, median, mode, standard deviation, and min and max for each day's height and yield.

The minimum, maximum, and plus and minus one standard deviation from the average was plotted using a box and whisker chart over time (days since planting) for yield, height, and yield per height. This was required to visualise the relationship instead of plotting each sample because the number of samples was larger than the maximum series per plot in the software we were using. A box and whisker plot is easier to interpret than a plot with a line for each sample and provides more information than a simple plot of the average over time.

The distribution was plotted for the yield, height, and yield per height for each day, and a summary for the entire dataset to show how the distributions change over time.
Since the same number of samples recorded for each day varied, the distributions were normalised so that plots for different days could be compared. The count of occurrences across the distribution were divided by the total number of samples in the distribution so that the y-axis showed the proportion of the total that each bin represents. Maximum values on the x-axis were chosen for yield, yield per height, and height to be appropriate to use the same scale for every day so that plots for different days could be compared. The minimum value for each x-axis was set at zero.

The scans were not analysed.

4.7.4 Procedure

The following steps were carried out to conduct the experiment.

4.7.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. Two trays were prepared using the standard method.
2. A template for seed placement was developed using a rectangular plastic lid with holes drilled about 10mm apart in a grid pattern of 5 by 7.
3. Seeds were placed on the coir using the template. Each tray had 210 seeds divided into 6 groups of 35 samples each, for a total of 420 samples. Each sample represented a single seed and were assigned a label (Figure 35).

4.7.4.2 Daily Procedure:

The following steps were completed for each day of the experiment.

11. The trays were watered as normal.
12. The height of each of the microgreen seedlings was measured and recorded (from 5 days after planting).
13. At least one section of seedlings was harvested (from 8-12 days after planting).
14. The total harvested part from each sprout was weighed.
15. The two cotyledons/leaves of each sprout were cut off and weighed.
16. The stems and leaves were taped flat to a blank sheet of A4 paper (using transparent tape), labelled, and scanned using a flatbed scanner.
17. For section L, digital vernier callipers were also used to measure the thickness of the leaves and stems.

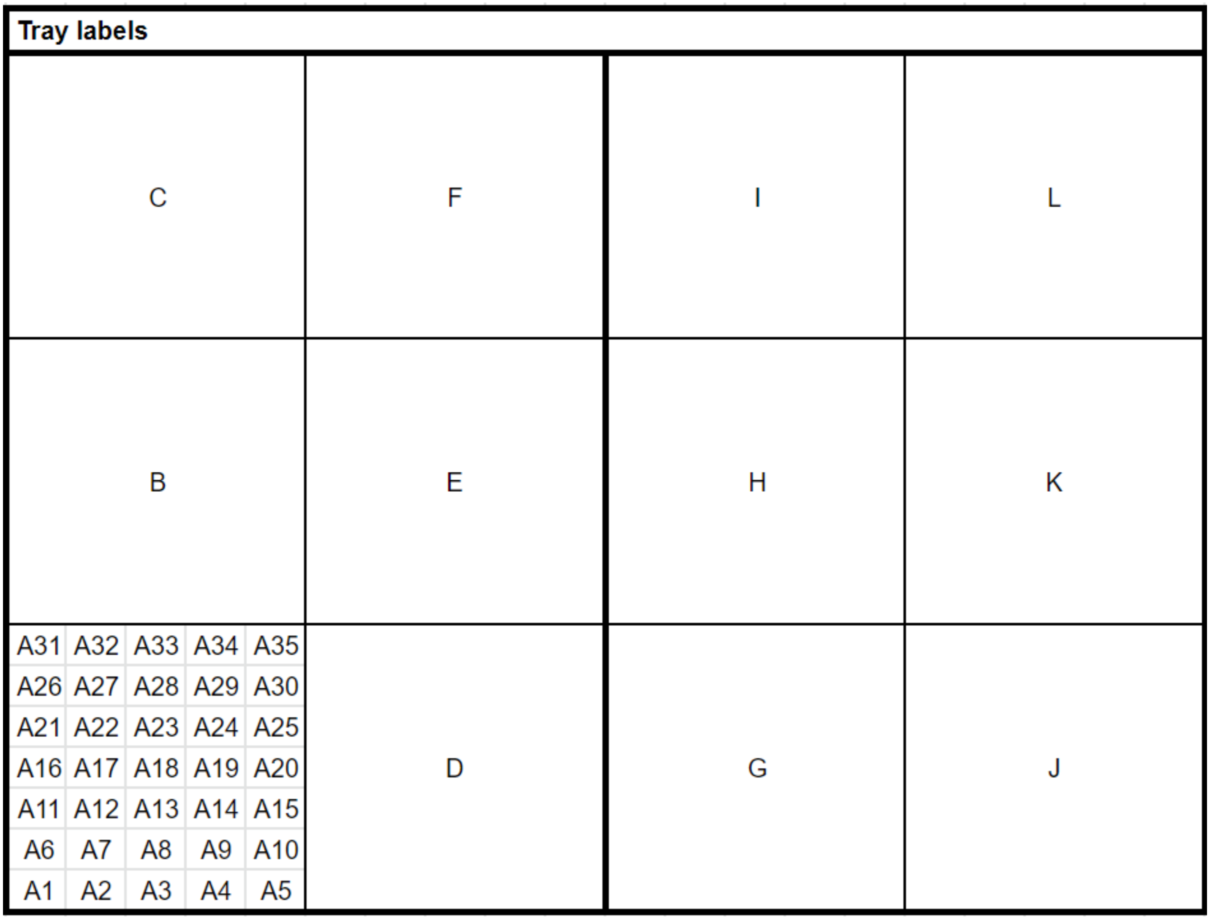

Figure 35 – Experiment 6: Diagram of the configuration of tray section and seedling labels

4.7.5 Results

For the box and whisker charts (candlestick charts) in these results, the bottom whisker of each candlestick represents the minimum value found across samples for that day, the top whisker shows the maximum value, and the body of the candlestick represents the average plus or minus 1 standard deviation. Charts were created to demonstrate yield over time per microgreen shoot, height over time per microgreen shoot, and yield per height over time per microgreen shoot (Figure 36, Figure 37, Figure 38).

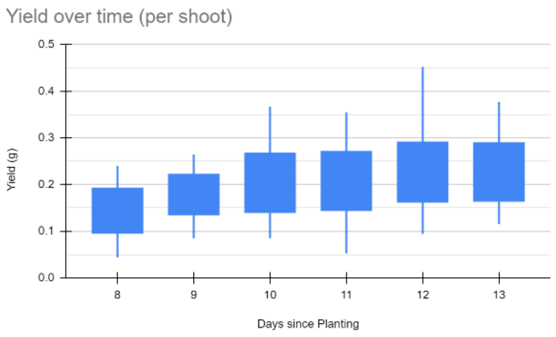

Figure 36 – Experiment 6: Candlestick chart summarising yield over time per microgreen shoot

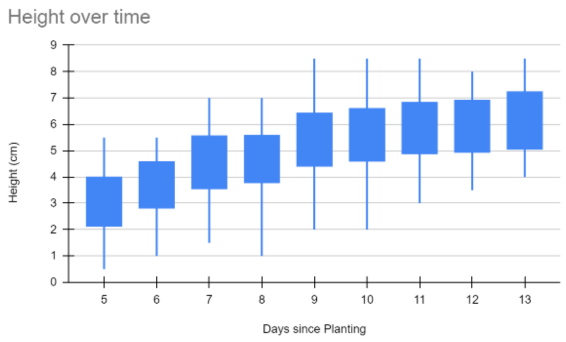

Figure 37 – Experiment 6: Candlestick chart summarising height over time per microgreen shoot

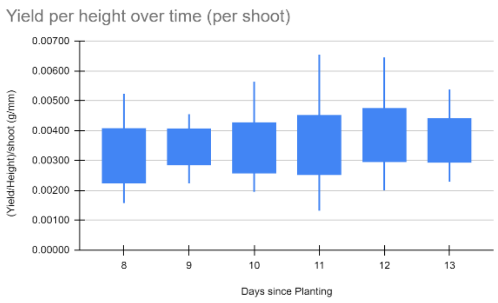

Figure 38 – Experiment 6: Candlestick chart summarising yield per height over time per microgreen shoot

4.7.6 Analysis and Conclusions

4.7.6.1 Challenges

- Some shoots moved from their original position, some did not grow, and many leaned over into the space occupied by others. This made it difficult to accurately keep track of which shoot belonged to which sample number. I recommend a change to this experiment where you use a 3D printed grid embedded in the soil to visually divide the shoots, and to keep them in position.
- Using tape to attach the samples to the paper worked, but not without problems. The tape didn't perfectly match the contour of the shoots, and it caused some undesired reflections and fingerprints. It is best to use a

minimal amount of tape. I recommend a change to using self-adhesive sticker paper instead of printer paper to eliminate the need for tape.

- There were times where we lost track of the correct sample label when attaching the shoots to the paper and had to spend time correcting mistakes. This was partly because some samples did not grow. Additionally, separating an image of each individual shoot was a tedious manual process because they were not consistently placed. I recommend printing a regular grid (with each square large enough to accommodate the largest sample). Each sample should be placed in its own square in order, leaving gaps for samples that didn't grow. The scanned images could then be divided into individual samples automatically using one of a few choices of free software.

4.8 Experiment 7

4.8.1 Goal

To contribute images to a database for computer vision analysis of microgreens with known heights, yields, and age, with fiducial markers present (coloured blocks of known size).

To use fiducial markers to automatically correct perspective for height estimation.

To determine whether watering prior to harvest would significantly affect the density or weight of the harvest.

4.8.2 General Method

Two trays of microgreens were grown using the standard method. A 3D printed block, and purchased connecting blocks were used as fiducial markers and were attached to the trays and to the wall behind the trays at a known. The trays were photographed each day and the markers were to be used to enable perspective correction.

At the conclusion of the experiment, the microgreens heights were measured, and they were harvested and weighed.

The harvest of each of the trays was place into water to determine the density.

Figure 39 - Experiment 7: Microgreens with coloured blocks as fiducial markers, Front

Figure 40 – Experiment 7: Microgreens with coloured blocks as fiducial markers, Side

4.8.3 Method of Analysis

The images from this experiment were not analysed because by the time the experiment was started, we had produced a python program for unrelated purposes that corrects image perspective based on ’ArUco’ tags. We abandoned this experiment in favour of using the program we had already produced, instead of spending time writing a new program to interpret these non-standard fiducial markers.

4.8.4 Procedure

The following steps were carried out to conduct the experiment.

4.8.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. Two trays were produced using the standard method.
2. A blue 3D printed block of known dimension was produced.
3. Black and orange plastic connecting blocks of measured dimensions were bought.

4. The black connecting blocks were joined in an ‘L’ configuration to consistently register against the corner of the rectangular trays.
5. An orange connecting block was attached to the top of the black ‘L’ so that it would always be in the same position relative to the tray whenever it is placed against the corner.
6. The blue 3D printed cube was secured at a height of 14cm at the base in the corner of a white wall. In this way, the blue cube should always be in the same position relative to the tray when the tray is pushed against the corner for a photo.

4.8.4.2 General Procedure:

The following steps were carried out on the day of harvest (day 8 after planting):

1. About half an hour before harvest, 300ml of water was poured into the bottom of one of the trays.
2. Prior to harvest, average height measurements were taken throughout each tray.
3. Photos were taken of each tray from all angles with the fiducial markers in place.
4. The harvest of each tray was weighed.
5. A large container was filled with 3000mls of water.
6. The harvest of each of the trays was placed into the water separately and the density and weight noted.

4.8.5 Analysis and Conclusions

The data from these experiments was not analysed. For future experiments, standard fiducial markers should be used with publicly available libraries. The goal of producing images with height and yield data for the database was achieved.

4.9 Experiment 8

4.9.1 Goal

To test whether the height and yield estimation algorithm is able to successfully estimate the height and yield with less than 20% error for radish microgreens grown under commercial conditions.

To contribute images of microgreens with known heights, yields, and age, and with ‘ArUco’ tags as fiducial markers to a database.

4.9.2 General Method

Radish microgreens were prepared using the standard for use commercially. Before harvesting for sale, images were taken from multiple angles with an ‘ArUco’ tag fixed squarely to the tray. These images were used to produce an estimate of the yield and height. The estimate was then compared against the actual measured yield and height.

4.9.3 Method of Analysis

To analyse the images from this experiment, they were passed as inputs to the yield estimation program. For more detail about how this works, see chapter 0

Algorithms and code used.

For tray 4, some minor modifications needed to be made to the program to analyse them, since the ‘ArUco’ tags were accidentally fixed to the tray at a 90-degree clockwise rotation.

The estimated heights and yields were compared against the measured heights and yields, and the errors were calculated.

4.9.4 Procedure

The following steps were carried out to conduct the experiment.

4.9.4.1 Set up:

The following describes the steps for the initial preparation of the experiment.

1. A clip was designed which would attach to the rectangular pots. It had three 40mm flat square orthogonal surfaces which would register in an external corner of the pots.
2. The 3D model was fabricated using a 3D printer.
3. Unique 40mm ‘ArUco’ tags were fixed to each flat surface using ‘blu-tak’ (Figure 41).
4. Trays of seeds were prepared using the standard method.
5. The clip with the tags was attached to the corner of the tray (Figure 42).

4.9.4.2 General Procedure:

The following procedure was carried out on the day of harvest (day 5 after planting):

1. Height was measured at multiple positions in each pot (as well as two days before harvest).
2. A photo (roughly) from the front, right side, and top was taken when the heights were measured (Figure 42, Figure 43).
3. The harvest yield was recorded.

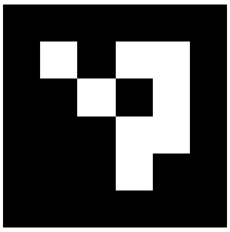

Figure 41 – Experiment 8: An example of a 4x4 'ArUco' tag with a marker ID of '0'

4.9.5 Results

The following images (Figure 42 to Figure 46) show examples of the input images, and the same images after the perspective has been corrected. The comparisons between the estimated and the measured heights and yields are tabulated in Figure 47.

Figure 42 – Experiment 8: Radish microgreens in tray with 'ArUco' tag affixed, Front (above)

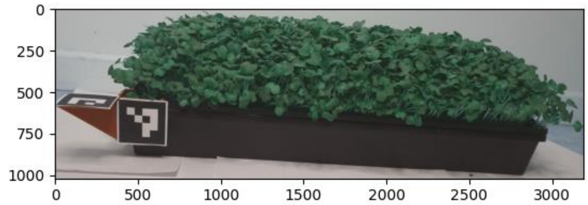

Figure 43 – Experiment 8: Radish microgreens in tray with 'ArUco' tag affixed, Side (above)

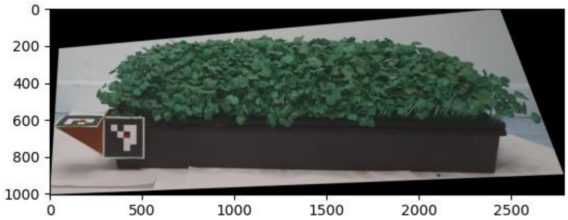

Figure 44 – Experiment 8: Radish microgreens in tray with 'ArUco' tag affixed, Perspective corrected from Side (above) to Side

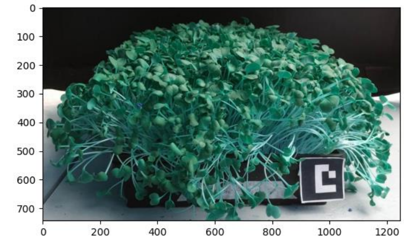

Figure 45 – Experiment 8: Example of radish microgreens input image

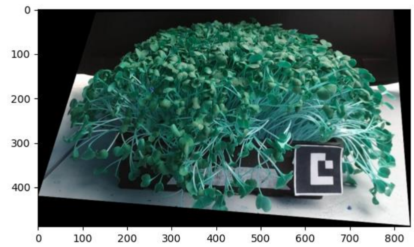

Figure 46 – Experiment 8: Example of perspective corrected image of radish microgreens

Trays:	Max heights measured around tray:				Age:	Measured values			Estimated values				
		Tray 1				Average height (mm)	Yield (g)	Est. Yield from height	Average height	Yield	Est. yield from height error	Height % error	Yield % e
25/9/22	110	115	115	100	8	104	518	681.5	103	811	31.56%	-1.40%	5
	100	115	115	100									
	100	120	115	90									
	100	90	90	90									
		Tray 2											
15/10/2022	105	110	110	Back	6	109							
	110	120	110	Middle									
	110	110	100	Front									
17/10/22	110	120	130	Back	8	124	521	815.0	94	753	56.42%	-24.32%	4
	120	130	130	Middle									
	130	120	130	Front									
		Tray 3											
20/10/22	70	65	55	Back	4	74							
Day 5	70	70	60	Middle									
	90	95	90	Front									
21/10/22	100	100	90		5	103		619.2					
	95	95	90										
	120	120	115										
23/10/22	115	110		Back	7	114							
	110	110		Middle									
	120	120		Front									
24/10/22	115	120		back	8	119	762	777.7	101	800	2.06%	-14.95%	
	120	120		middle									
	Front half weight 402g												
	Back half weight 360g												
		Tray 4											
31/10/22	130	125	120	Back	8	129	710	847.7	65	527	19.39%	-49.43%	-2
	140	140	130	Middle									
	125	130	125	Front									

Figure 47 – Experiment 8: Table of data comparing estimated and measured heights and yields, showing error of up to 57%

4.9.6 Analysis and Conclusions

While the yield and height estimation program was only able to be tested on four samples because of time constraints, the results were clear.

Only two of the four samples were estimated to within ±20% for height, and only one for yield. The maximum error was nearly 57%. Further, the error in height and yield estimations did not correlate strongly with each other. This indicates that the height and yield for the samples were not strongly correlated themselves. Finally, the absolute errors were large enough that the difference in the yield per height due to the age of the shoots was insignificant.

A challenge with this experiment was difficulty detecting the ‘ArUco’ tags. Because the tags were printed and cut to the edge of the tag, but the tags were also in front of a black backdrop (the trays), there was not enough contrast to consistently detect the tags. Occlusion due to overhanging leaves made this problem worse. Because the tags were printed on paper, they were bent if they were not securely fastened to a firm, flat backing. An improvement to the method for this experiment would be to print the ‘ArUco’ tags smaller, allowing room for white-space around them, and on stiff cardstock. Printing on the same card twice may also help increase the contrast between the black and white. Of course, care should also be taken to ensure the tags are not occluded.

Goal achievements:

Overall, the goal of testing the functionality of the system was achieved, even though the result was that it was not functional.

The goal of making contributions to the image database for further research was partially achieved. Some contribution was made, but the contribution was small (images for four samples), and there were errors in some of the images (‘ArUco’ tags being placed incorrectly).

5 Algorithms and code used

We used python to implement the yield estimation algorithm.

The external libraries used were cv2, matplotlib pyplot, numpy, and rawpy (and their dependencies).

The main function for the program that we wrote to estimate the yield is included (Figure 48 to Figure 55) and explained below. The full code that is required for the yield estimation, including the (well commented) functions that we wrote that are used by 'main' have been included in Appendix B: Developed functions, algorithms, and code.

```python
def main():
    # Input parameters
    A_mm2 = 255 * 515     # Total area of the pot
    h_offset = 15           # Distance between the soil and the top of the pot
    age = 10                # Age in days
    image_path = "./images/"
    image_filename = "image3.jpg"

```

Figure 48 – Code: Main function, input parameters

```python
    # inits
    aruco_tag_size_mm = 49
    aruco_tag_size_pxls = 100
    w_mm = 255              # Pot width (mm) (measured)
    d_mm = 515              # Pot depth (mm) (measured)
    h_tray_mm = 70          # Pot height (mm)
    A_mm2 = w_mm * d_mm     # Total area of the pot

    seed_grams_per_mm2 = 55/(255 * 515)  # Grams per mm^2 (55 grams in a standard pot)
    seed_count_per_g = 690/11            # Seed count/gram (Determined experimentally)
    seed_viability = (420 - 33)/420      # Viable seeds percentage (Determined experimentally)
    num_shoots = A_mm2 * seed_grams_per_mm2 * seed_count_per_g * seed_viability
    num_splits = 5                      # More splits gives more samples for an average
    h_mm_per_pxl = aruco_tag_size_mm/aruco_tag_size_pxls    # Height in mm per height in pixel
```

Figure 49 – Code: Main function, variable initialisations

```python
    # Load image (jpg or raw)
    try:    img = rawpy.imread(image_path+image_filename).postprocess
    except: img = cv2.imread(image_path+image_filename)
    plt.imshow(img)
    plt.show()
```

Figure 50 – Code: Main function, loading image

```python
    # Warp image to correct for perspective
    warped_img = manipulate_tags(img, aruco_tag_size_pxls
    plt.imshow(warped_img)
    plt.show()

    img = warped_img
```

Figure 51 – Code: Main function, perspective correction

```python
    # Convert image to HSV colour space
    img_hsv = cv2.cvtColor(img, cv2.COLOR_RGB2HSV)

    # Perform image cropping/masking to keep only greens
    img = hsv_crop(img_hsv)
    plt.imshow(img)
    plt.show()
```

Figure 52: Code – Main function, HSV colour cropping/segmentation

```python
    # Remove noise and fill in holes - all pixels left should be part of pla
    img = denoise(img)
    plt.imshow(img)
    plt.show()
```

Figure 53: Code – Main function, removing image noise

```python
    # Measure average maximum height in pixels
    h_pixels, w_pixels = get_pxl_extent(img, num_splits)

    # Convert pixel height to mm
    h_mm = h_mm_per_pxl * h_pixels
    average_max_height = h_mm + h_offset

    average_height = average_max_height/2
```

Figure 54: Main function, height estimation

```python
    # Get yield estimate based on height (mm), age (days), and estimated shoot cou
    yield_estimate = get_yield_for_height_at_age(average_height, age, num_shoots)

    print("num_shoots", num_shoots)
    print("h_pixels", h_pixels)
    print("w_pixels", w_pixels)
    print("h_mm", h_mm)
    print("yield_estimate", yield_estimate)
    print("-----------------------------")
```

Figure 55: Code – Main function, yield estimation and printing outputs

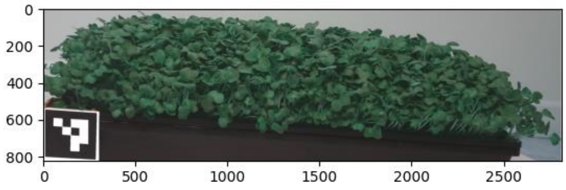

Figure 56: Code example image 1, input image

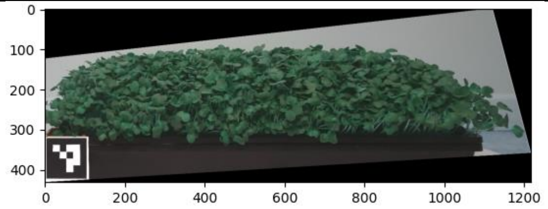

Figure 57: Code example image 2, perspective correction applied based on fiducial marker

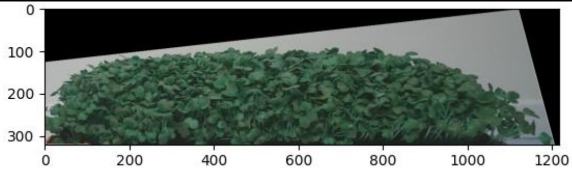

Figure 58: Code example image 3, image cropped based on fiducial marker

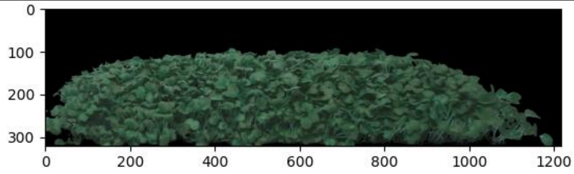

Figure 59: Code example image 4, microgreens segmented based on HSV colour thresholds

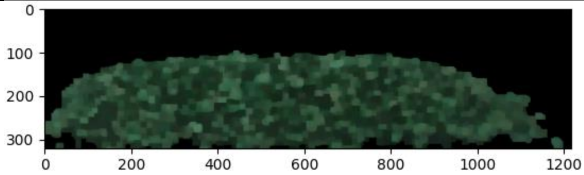

Figure 60: Code example image 5, denoising function applied to segmented microgreens

```
num_shoots 3178.928571428571
h_pixels 209.4
w_pixels 1198
h_mm 102.606
yield_estimate 811.2785296499999
------------------------------
```

Figure 61: Code example image 6, example output for main function

6 Overall Results

We plotted the yield, height, and yield per height relationship, as measured for individual microgreens. We found, as expected, that both the height and the yield increased over time, with the rate of change decreasing as the shoots got older. The height and yield across the data had a high degree of variability. We found that the yield per height relationship was approximately linear, with grams per mm per shoot (with age in days since planting).

This relationship only held with individual shoots and could not be used to estimate the average yield of entire populations. When applied to estimate yields of trays of microgreens with known heights, the error varied from 2% to 56%. We developed an algorithm to estimate the height of microgreens and then applied the yield-height relationship to estimate the yield. When tested against trays of microgreens grown for commercial purposes, the algorithm's accuracy varied between 1.4% and 56.6%, which was below the minimum required precision for it to be useful (20%).

I reviewed the findings with the industry partner. Although the minimum required precision of 20% was the target, the industry partner found the variance of up to 56.6% interesting and encouraging. The range of results meant that there could be a lot of potential for increasing average yields through understanding, controlling, and "tweaking" the variables involved in microgreens production.

We produced over 1000 images of microgreens (along with the corresponding yield and height data) during the experiment

7 Overall Analysis and Conclusions

7.1 Summary

The project was conducted well and completed on time. We executed according to our planned approach, making several adjustments along the way.

Although we did not achieve the goal of 20% yield estimation accuracy, our approach and learnings will inform subsequent work well, and positions the industry partner well for next steps.

This section details the projects problems, improvements and achievements met against goals.

7.2 Problems encountered

The yield and height estimations errors were far larger than required for useful results. There are a few problems and observations that could explain this.

1. We may have an undetected error in the code or a logical error in the algorithm.
2. We evaluated the algorithm against plants grown for commercial use, which we grew later in the year than the samples used to establish the yield-height relationship. Slight differences in experimental conditions or procedure might result in significant differences in their growth. Since there are many more trays to maintain in a commercial system, an individual tray might get different care and attention as a tray being carefully grown for experimentation. This may result in lower or more varied yields and heights.
3. One of the major difficulties with experiments on plants is that plant growth can be affected by various environmental factors, which can interact in complex and non-linear ways and be difficult to control or sometimes even measure. For example, the temperature, humidity, airflow and soil aeration and retention, and light intensity and frequency are just some of the major

factors that influence the water requirements of a plant. Most of these factors constantly change over time with limited predictability. The plant's water requirements (and how adequately they are being met) are just one of the many factors influencing the plant's growth rate. We controlled the soil-related parameters for these experiments by using the same growth medium across experiments (except for experiment 1). The light intensity and frequency, the air flow, and the temperature (to a limited extent) was controlled by keeping the samples in an outdoor structure with fans and time-controlled grow lamps. Because we could not perfectly control these conditions, the required water between samples grown at differed with time, and how well those needs were met varied. This resulted in different growth profiles for plants grown even a couple of months apart.

4. Since we harvested and measured the shoots in experiment six individually, we likely took more care to harvest close to the soil than when harvesting an entire tray. However, we expect the height measurements to be consistent since we used the soil level as a reference to register the ruler. We expect that this would result in an overestimation of the yield. This is (mostly) what we see. However, it does not explain the variability in the error.

5. Experiment six found a high degree of variability in individual shoots' yield per height relationship. We assumed that due to the law of large numbers, using an average height across a population would result in a more accurate estimation for the population than estimations of individual shoots. However, there was also a large degree of variability in the estimated yields and average heights of the trays of microgreens. It is possible that this is because of inaccuracy in the height estimation system. For example, when the perspective is corrected to warp the image into a ‘front on’ or ‘side on’ view, this process may not decrease the perceived height of the microgreens to the same as it would be from an image taken straight on, as it was assumed it would. This would result in an overestimation that depends on the perspective of the original input image and may appear random. This is exactly what we see.

6. Another potential cause for seemingly random overestimation is if the opening and closing operations (intended to remove noise) does not remove all of the pixels outside of the segmented plants. Because of the way the height estimation works, a single stray pixel could significantly increase the height and yield estimates.
7. In experiment 6, we found that the yield contribution due to the stems and the yield contribution due to the leaves was only weakly correlated for those samples. In experiment 8, we applied the yield-per-height relationship to known heights to evaluate the yield-height relationship without estimating the height. We found that there was still a lot of variability in the error. These results could mean that the leaf and height yield contributions are even less strongly correlated in the commercial examples.

While stem and leaf yields not being strongly correlated means that measuring the height alone likely will not provide enough information to estimate the yield accurately, it is also useful information. The industry partner has stated that this means they can improve some microgreens' characteristics without significantly affecting the yield. For example, leaving microgreens in darkness for longer causes them to grow taller, which makes them easier to harvest. If there is little correlation between the height and the leaf yield, then this means that this might be able to be done without it decreasing the yield from the leaves.

7.3 Suggested improvements

These numbered suggestions for improvements correspond to the numbered problems encountered in section 7.1. These suggestions describe ways to improve this work to allow more confidence in analysis and results, and to improve the systems "as-is". Suggestions for other research projects based on this work, or with the same broad goals as this one can be found in chapter 8.

1. It can sometimes be difficult to detect errors in your own code or logic without an outside opinion or a sounding board. Anyone attempting to duplicate this study should first invest in analysing whether any assumptions or logical

steps that we made were well founded. Scrutiny should be addressed to our methods of using height and yield measurements of individual sprouts, and how this relates to measurements and estimates of trays of sprouts, as well as the code.

2. To improve the consistency between laboratory experiment measurements and real-world measurements, we can set up standards to follow between them so that the results can be compared. Since the business methods should not be changed (unless the change will improve profits), the experimental procedure should match the real-world commercial procedure. Including the variability in the samples that this might cause.

3. For the issue of controlling and measuring the many plant growth parameters, there are two opposing pathways for this improvement.

    a. Drastically improve the level of control over environmental factors. Perform experiments in a sealed environment where all inputs and outputs are controlled and measured. Perform the experiments where there is a large thermal mass, or underground, to allow finer control over the temperature. For this data collection stage, depth cameras and all manner of precise equipment can be used to monitor and control the experiment. Perform many experiments under many different iterations of environmental conditions until their relationships are well understood and documented. Then use this new precise model of plant growth in an (uncontrolled) real world system using sensors to estimate yield.
    This approach has the advantage of improving knowledge of plant growth with a precise model. However, it would be expensive and take a long time.

    b. Acknowledging that the practical goal is to be able to apply the model to crops frown under real-world conditions, we can collect height and yield data from crops grown for commercial purposes instead of instead of ‘training’ the height-yield-age relationship using individual shoots grown under laboratory conditions. This has the advantage of all the measurements corresponding to real-world conditions. It also has the advantage that there are no costs associated with the materials in the experiment (since they are being sold anyway). The disadvantages are that it may take a long time to collect enough data for underperforming samples to predict them, since the goal of commercially grown plants is generally to minimise underperformance. Also, experimental procedures would need to adhere to the appropriate standards for food handling.

4. The solution to improve consistency between cuts is similar to that for other measurements, with one difference. It is possible to use (or develop, if required), a system to automatically (and consistently) harvest trays. If this is beneficial for the business, it can be used for experiments and for commercial use.

5. The error due to perspective was attempted to be corrected, but the accuracy or applicability of this correction was never evaluated. This correction was implemented so that an image could be taken with a phone under non-ideal conditions, in the real world without training or special equipment. However, it is now clear that it would have been prudent to prove that the system works under ideal conditions before attempting to adapt it to non-ideal conditions. The error in the height estimation due to perspective can easily be nearly eliminated in laboratory conditions. If the photo is taken from a distance with a camera with a zoom lens, centred on the tray, the resulting image will have minimised perspective distortion.

6. To test the opening and closing noise removal step of the height estimation algorithm, a number of estimated samples can be compared against the same samples after being manually segmented. If they are significantly different, then this is an issue. This issue can be fixed by implementing further noise removal steps.

7. If the reason for the variability is that the leaf and stem yields are not strongly correlated, then the model of using the yield-height relationship to estimate yield will not work on its own. A more complex model will need to be

developed that separately considers the contribution due to the leaves and stems. There are more suggestions for this in chapter 8.

7.4 Goal Achievements

The Specific Goals for the Thesis are restated here for convenience.

1. Create a model that estimates yield (in grams) of microgreens with an error of at most 20% from images or video captured by a modern smart phone device. The industry partner commissioning this study estimates that this would be sufficient to have useful merit for their business as a first investigation.

   Result:

   The model developed was not able to estimate yield of microgreens with the required precision to be useful.

2. Evaluate and record extensions and limitations of the model and identify a path for improvement for future work.

   Result:

   We did not evaluate extensions of the model due to time constraints, and because the basic use of the model was not successful.
   We succeeded in identifying limitations of the model, and have laid out a suggested path for improvement.

3. Start a database of images and associated data that could be used for future study.

   Result:

> We produced a database of over 1000 images (and their associated height and yield measurements). The image set includes photos of radish microgreens under various lighting conditions, from various angles, and with various fiducial markers. The image set also includes scans of individual shoots.

Overall, we achieved the stated specific goals, though the contribution of this work towards the general goal of increasing accessibility of experimentation on plants for small businesses and individuals was minimal.

8 Recommendations for future study

The thesis findings demonstrate the opportunity for extensive further study. Additional studies could significantly impact the efficiency of small growers both in the provision of insights from the findings as well as tools that could result from development.

We have provided some suggestions for further study in this section.

8.1 Collecting more data to improve results:

Repeating experiment 6 to collect more height-yield data for individual shoots may result in better accuracy for the height-yield relationship and more detailed knowledge of their height and yield distributions.

Experiment 8 suffered from time constraints causing only a few data points to be collected.

The time constraints limited the ability to test the system rigorously. Collecting more height and yield data, along with images of the corresponding microgreens with 'ArUco' tag fiducial markers, would allow for better analysis of the system's accuracy.

Further, with a large enough number of these samples (with known yields), a yield-height relationship can be established using the heights that the program estimates from the images rather than individual shoots. The accuracy of the two results for the yield-height relation could then be compared.

8.2 Expansion of parameters being measured or controlled:

The height-yield-time relationship was explored for a single growth period (between 1 and 2 weeks). But there are many factors that we still need to explore in this work that significantly affect plants' growth. For example, the yield-height relationship over more extended periods (e.g., over a year) is unexplored by these experiments. Environmental factors that change throughout the year (mainly temperature, rather than sunlight, since these samples were grown under grow lamps) significantly affect plants' growth.

The company's commercially grown radish microgreens can be harvested from 8 days after planting in the winter but 6 days after planting in the spring due to an increased growth rate. As a result, the yield-height relationship we used in this study (which we determined from samples grown primarily in July and August in southeast Queensland, Australia) may not apply to plants grown during other times of the year or in different climates.

For a future study, this experiment could be repeated throughout the year or in different climates to explore how these affect the yield-height relationship.

Factors such as temperature, humidity, and soil moisture can be measured to help draw valuable conclusions.

General weather information can be obtained using public weather data (e.g., from bom.gov.au), but it would be better (more accurate) to use data captured continuously using sensors placed in the microclimate of the experiment.

For collecting sensor data, we recommend the open source 'b-parasite' soil moisture, ambient temperature, and humidity sensors to collect this data [11] (Figure 62). They must be manufactured from the freely available designs since they are not sold as a complete product. Based on the versatile ESP32 microprocessor, these devices are relatively inexpensive to produce. They can broadcast via 'Bluetooth Low Energy' and are expected to last up to a year powered by a coin cell battery. They natively support the free sensor and actuator software platform 'ESPhome' , making them easy to implement in a DIY monitoring system. We manufactured several for our experiments but could not implement them in the allowed time.

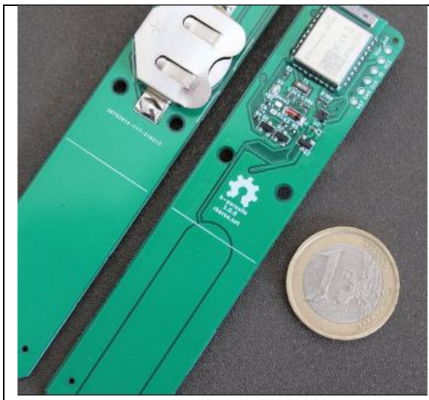

Figure 62 – Image of a 'b-parasite' sensor, Front and Back. Australian $1 coin for scale.

There are many other parameters that could be varied to explore their effect on the height-yield relationship.

For example, the effect on the yield-height-time relationship caused by:

- Humidity, air pressure or gas composition.
- Time and quantity of watering.
- Time and intensity of light under grow lamps, as well as the timing throughout the day-night cycle when the lamps are turned on and off.
- Density and pattern of seed spread, including random placement vs organised grids.
- Variation in the brand, batch, or age of the radish seeds.

8.3 Investigation of other microgreen species:

We chose radish microgreens from a variety available to use, partly because of the small variation between the lowest and highest leaf height, which we assumed would lead to a better estimate of the average height of the leaves. Radish microgreens also have the feature that the colour of the stems compared to the leaves is easily distinguished (which allows the stems to be digitally removed to allow for the average leaf height to be calculated). Future experimentation is needed to establish the yield-height relationship and determine if this yield estimation method is valid for other species.

For future studies, our first recommendation of a species to use is broccoli because of its similar height and colour characteristics to radish (Figure 63), and because it is a popular product . [4]

Our next recommendation is sunflower. Sunflower microgreens have greater variation in height than radish, and their stems and leaves are similar in hue (i.e., hard to segment) (Figure 64). They have a greater overall height and are also a popular product with versatile uses. The increased overall height at maturity might benefit the height estimation's accuracy.

Other examples of microgreens that could be investigated include Asian greens, beetroot, buckwheat, carrot, cos lettuce, cress, kale, mesclun, peas, rocket, silver-beet, tatsoi, turnip, and wheat.

Figure 63 – Image of broccoli microgreens, showing consistent height growth between sprouts, and the colour difference between stems and leaves.

Figure 64 – Image of young sunflower microgreens, showing variation in height growth between sprouts, and the similarity in colour between stems and leaves.

8.4 Improving library of images for computer vision research:

The database of photographs and corresponding measurements already established in this work could be used for other plant-based computer vision projects. Many of the datasets currently used for image-based plant phenotyping analysis and automated leaf segmentation use only images of tabaco or tomato plants. The dataset of images from this work offers alternatives. The dataset contains images of microgreens and their yields and heights in various conditions. For example, cluttered and uncluttered environments, various lighting conditions, and examples with and without a variety of fiducial markers (‘ArUco’ tags, rulers, and coloured cubes). The dataset also includes scans from a flatbed scanner of the microgreens, and most of the images are in both raw and jpeg formats. This variety in the dataset enables broad possibilities for further computer vision ‘proof of concept’ studies. Researchers could use it to test techniques and refine their own experimental setups before committing to collecting their own data.

The drawback to this dataset is that there may not be enough data of any one type (e.g., images under certain lighting conditions, at certain angles or with certain fiducial markers) to enable a large-scale research project.

Further research could be done to enlarge or otherwise improve this dataset.

For example, this study would benefit from more images of microgreens with ‘ArUco’ tag fiducial markers.

8.5 Improving accuracy with more complex physical models:

The current system uses a model of only how height affects yield. Since the stems primarily contribute to the height and the leaves mainly grow horizontally, the system's accuracy is limited by the strength of the correlation between the contribution to the yield of the stems and the leaves. Therefore, a more complex model that considers the leaves could increase the prediction's accuracy.

Since microgreens are grown densely packed, many occlusions (overlapping) of leaves exist. Because of this, the size of the canopy area needs to provide more information about the yield. Also, leaf segmentation is a difficult task for microgreens because of this occlusion. Additionally, since microgreens don't have true leaves, counting the number of leaves does not provide any information (since a healthy dichotomous plant without true leaves will always have two leaves). For true leaves, the midribs and veins of the leaves can be used to help fit a leaf model, but the veins of cotyledons are often more challenging to distinguish. However, there may be a way to estimate the average size of the leaves.

Using the existing scans and yield data from experiment 6 of this thesis, a leaf area to leaf yield relationship can be determined. Using these scans again, deformable models of ‘prototype’ leaves could be developed (Figure 65). If the partially occluded leaves can be segmented, prototypical leaf model could be scaled and deformed to fit the unoccluded parts of each segmented leaf. Once a best-match fit is found for each leaf, the area of the digital prototype leaves can be used to estimate the actual leaf area (Figure 66). Using this estimate for the leaf area, we can calculate an estimate for the contribution to the yield due to the leaves. The contribution to the yield due to the stems can be estimated using the current method to estimate the total yield, scaled down by the proportion of the stem yield to the total yield (which is known from experiment 6).

Figure 65 – Example of a scan of a microgreen leaf which could be used to develop a digital ‘prototype’ leaf model.

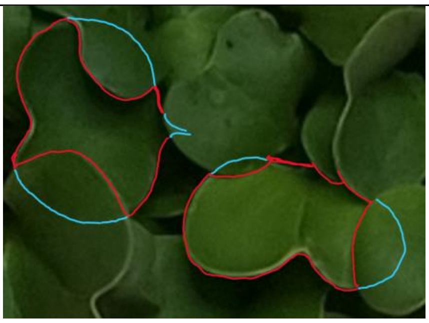

Figure 66 – Example of how an occluded leaf might be segmented (red), and how fitting a deformed ‘prototype’ leaf model might correct the leaf’s calculated area.

8.6 A modular system for conducting experiments:

One of the problems this work is trying to solve is that consistently collecting and analysing plant data is difficult. But we can make further collection easier by collecting and analysing enough data correctly.

Specifically, in this thesis, we collected height and yield data. We then used that data to develop a method to automatically collect height and yield data. However, this only applies to radish microgreens grown in July-August under the same environmental conditions. Collecting the required data for a similar analysis for any variation would require a similar effort.

Developing a modular system for conducting these experiments would decrease the work required for each variation and enable greater consistency between experiments. Each module should contain its own variable intensity grow lights and an environmental condition sensor suite (temperature, soil moisture, etc) that can be monitored and controlled from a central online location. Each module should also have a feature that facilitates images to be taken consistently and under even lighting. The module should be cheap enough to be accessible and practical for small businesses and individuals ('citizen scientists'). There may be multiple models designed. For example, there may be one model designed to facilitate experimentation and a cheaper model with fewer features designed to be practical for monitoring crops for production.

The system should also be modular and scalable to be useful to individual growers and small businesses.

An app could also be developed to streamline inputting and interacting with the data.

8.7 Performance prediction:

By collecting the yield-height growth data over time for many samples, a baseline can be established for how a sample is likely to perform in the future, depending on its current age and height. The data used will need to include overperforming and underperforming samples.

One complicating factor that may need to be overcome is that the reason a crop is underperforming may affect the outcome. For example, consider two crops; both underperformed by 15% at day 5. One is underperforming due to overwatering, and the other is due to underwatering. Let's say (arbitrarily for this example) the underwatered crop underperforms by 15% at day 8, but the overwatered crop underperforms by 18% at day 8. If we only consider the current performance and age, there is no way this could be predicted at day 5.

One way that these modes might be distinguishable is by considering the past height, and/or the profile of the rate of change of the height over time. For example, (arbitrarily for this example), the overwatered crop might grow faster than the underwatered crop for the first few days, and then the growth rate might decrease.

On the other hand, the underwatered crop might have a growth rate that is mostly linear (after normalising by the expected growth over time for an average sample).

8.8 AI-based approaches:

We used an analytical method for determining the yield – defining a one-dimensional relationship between yield per height and time. There are too many methods of artificial intelligence to go into detail here, except to say that systems such as ‘convolutional neural networks’ (CNN) and ‘deep learning’ or ‘deep convolutional neural networks (DL or DCNN) are examples of general AI methods that could be applied to the problem of gaining insights (such as yield estimates) using images of plants.

CNN and DL algorithms can consider relationships between information in the image without this being explicitly programmed. For example, the analytical method we used considers only the height estimated using pixels segmented based on simple colour thresholds. It then uses a static yield-height relationship to estimate yield. A CNN or DL might draw a relationship between the colours of the leaves and stems and use that to modify the yield-height estimation. The AI-based system could then be compared against our solution for accuracy.

Other insights that future research could investigate using AI may include detection of pests, wilting, and mould.

It should be noted that deep convolution neural networks often require millions of data points.

9 References

[1 J. Lee, “Determination of Leaf Color and Health State of Lettuce using Machine Vision,” *Journal of Biosystems Engineering,* vol. 32, no. 4, pp. 256-262, 2007.

[2 S. D. Khirade and A. Patil, “Plant Disease Detection Using Image Processing,” *IEEE,* 2015.

[3 R. L. C. Z. X. Q. C. D. C. S. H. L. Rujing Wang, “An automatic system for pest recognition and forecasting,” *Pest Management Science,* vol. 78, no. 2, pp. 711-721, 21 October 2021.

[4 C. o. NaturalYield, Interviewee, *Discussion with Natural Yield Pty Ltd.* [Interview]. March 2022.

[5 M. J. J. A. a. G. K. T. Mueller-Sim, “The Robotanist: A ground-based agricultural robot for high-throughput crop phenotyping,” *2017 IEEE International Conference on Robotics and Automation (ICRA),* 2017.

[6 Thermometer World, *Thermometer World bringing measurement to life,* 2022.

[7 Dr Meter, *Dr.meter,* 2022.

[8 The IOT Store, *DFROBOT GRAVITY: ANALOG AMBIENT LIGHT SENSOR FOR ARDUINO,* 2022.

[9 Applied Drone Technology, “Precision Aerial Mapping,” 2020. [Online]. Available: https://www.applieddronetechnology.com.au/precision-aerial-mapping.

[1 E. S. K. A. F. T.-R. W. P. K. H. S. N. Emre Tunca, “Estimation of bell pepper evapotranspiration using two-source energy balance model based on high-resolution thermal and visible imagery from unmanned aerial vehicles,” *Journal of Applied Remote Sensing 16(2),* 2022.

[1 Raphael, “b-parasite,” [Online]. Available: https://hackaday.io/project/178651-b-parasite#:~:text=Description%20b-parasite%20is%20an%20open%20source%20soil%20moisture,it%20via%20Bluetooth%20Low%20Energy%20%28BLE%29%20advertisement%20packets..

[1 Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra., “ABARES Insights 2020,” *Analysis of Australian Food Security and the Covid-19 Pandemic,* p. 4, 17 April 2020.

[1 FAO, IFAD, UNICEF, WFP, WHO, “THE STATE OF FOOD SECURITY AND NUTRITION IN THE WORLD 2021,” in *Transforming Food Systems for Food Security, Improved Nutrition and Affordable Healthy Diets for All*, Rome, FAO, 2021, pp. 1-240.

[1 M. Bowden, “Understanding food insecurity in Australia.,” Child Family Community Australia, Southbank, 2020.

[1 United Nations, DESA, Population Division, “World Population Prospects 2022,” 2022. [Online]. Available: https://population.un.org/wpp/Graphs/Probabilistic/POP/TOT/900.

[1 D. Marin-Guzman, “Supply chain crunched as up to half of truckies isolate,” 5 January 2022. [Online]. Available: https://www.afr.com/work-and-careers/workplace/supply-chain-crunch-as-up-to-half-of-truckies-isolating-20220105-p59m22.

[1 The IOT Store, *ANALOG CAPACITIVE SOIL MOISTURE SENSOR-CORROSION RESISTANT,* 2022.

[1 The IOT Store, *SHT20 I2C TEMPERATURE & HUMIDITY SENSOR (WATERPROOF PROBE),* 2022.

[1 Pakronics, *Gravity: Analog Water Pressure Sensor,* 2022.

[2 InstrumentChoice, *MK35ON Premium Handheld Spectrometer.*

10 Appendices

10.1 Appendix A: DevOps Task list

This was the initial plan produced at the start of the thesis research, consisting of tasks that we expected we would need to complete.

Work Item Type	Title	State
Epic	Thesis Admin	Active
Feature	Submit Thesis Proposal	New
Feature	Seminar Presentation	Active
User Story	Prepare	New
User Story	Dry Run	New
User Story	Execution	New
User Story	Participate in 5 other seminars	New
User Story	Online academic integrity tutorial	Closed
Feature	Poster & Demonstrations	New
User Story	Create poster and present for demonstration	New
Epic	Project Proposal	Active
Feature	Research	Active
User Story	Define scope of research	Active
Task	Reading of the subject / domain	Active
Task	Identify an issue to solve	Active
Task	Identify gaps in the product space / commercial ma...	Active
Task	Plant species	Active
Task	Length of testing cycles	Active
Task	Infrastructure design (equipment racks)	Active
Task	Current state-of-the-art research	Active
Task	Identify gaps or limitations in the state-of-the-art	Active
User Story	Create Project Proposal	Active
Task	Identify problem or issue to solve	Closed
Task	Define general goals	Closed
Task	Get feedback from end client	Closed
Task	Identify current technologies / techniques	Closed
Task	Identify prior work	Active
Task	Identify state-of-the-art in prior work	Active
Task	Identify gaps or limitations in the prior work	Active
Task	Identify how my research could fill the gap	Active
Task	Risk assessment	Closed
Task	Breakdown of all of the planning	Active
Feature	Feedback	New
User Story	Weekly supervisor meetings	Active
User Story	Submit Project Proposal	New
Feature	DevOps	New
User Story	Create DevOps Work Breakdown	Active
User Story	Create Sprints	Active
User Story	Review and Revise Plan	New

Work Item Type	Title	State
Epic	Stage 1 - Prototype using image analysis	New
Feature	Identify & Define	New
User Story	Research different tools for consideration	New
User Story	Choose and implement tool(s)	New
User Story	Learn Tools	New
User Story	Design prototype infrastructure	New
Feature	Implementation Setup	New
User Story	Purchase infrastructure / equipment	New
User Story	Build prototype infrastructure	New
User Story	Test infrastructure / equipment	New
Feature	Implementation Cycle 1	New
User Story	Planting seeds	New
User Story	Daily maintenance of equipment	New
User Story	Daily observations of plants	New
Feature	Inspect and Adapt Cycle 1	New
User Story	Infrastructure and Equipment adjustments	New
User Story	Review and adjust experimental data collection proced...	New
User Story	Modify and improve (as needed) data processing tech...	New
Feature	Implementation Cycle 2	New
User Story	Planting seeds	New
User Story	Daily maintenance of equipment	New
User Story	Daily observations of plants	New
Feature	Inspect and Adapt Cycle 2	New
User Story	Infrastructure and Equipment adjustments	New
User Story	Review and adjust experimental data collection proced...	New
User Story	Modify and improve (as needed) data processing tech...	New
Feature	Stage 1 Result Analysis	New
User Story	Compile data for presentation	New
User Story	Compare and assess predicted vs measured outputs (...	New
User Story	Identify input-output correlations	New
User Story	Transform collected images to quantitative data	New

Work Item Type	Title	State
Epic	Stage 2 - Using advanced computer vision techniques	New
Feature	Identify & Define	New
User Story	Review and revise design for tools / setup	New
User Story	Update Tools	New
User Story	Design updated infrastructure	New
Feature	Training Model (using stage 1 data)	New
Feature	Implementation Setup	New
User Story	Purchase infrastructure / equipment	New
User Story	Build prototype infrastructure	New
User Story	Test infrastructure / equipment	New
Feature	Implementation Cycle 1	New
User Story	Planting seeds	New
User Story	Daily maintenance of equipment	New
User Story	Daily observations of plants	New
Feature	Inspect and Adapt Cycle 1	New
User Story	Infrastructure and Equipment adjustments	New
User Story	Review and adjust experimental data collection proced...	New
User Story	Modify and improve (as needed) data processing tech...	New
Feature	Implementation Cycle 2	New
User Story	Planting seeds	New
User Story	Daily maintenance of equipment	New
User Story	Daily observations of plants	New
Feature	Inspect and Adapt Cycle 2	New
User Story	Infrastructure and Equipment adjustments	New
User Story	Review and adjust experimental data collection proced...	New
User Story	Modify and improve (as needed) data processing tech...	New
Feature	Stage 2 Result Analysis	New
User Story	Compile data for presentation	New
User Story	Compare and assess predicted vs measured outputs (...	New
User Story	Identify input-output correlations	New
Feature	Forward testing with trained data	New
Epic	Stage 3 - Generalize model from stage 2	New
Feature	Extend and Refine Stage 2 Tools	New
Feature	Implementation	New
Feature	Inspect and Adapt	New
Epic	Thesis Report Submission	New
Feature	Draft Final Thesis	New
Feature	Submit Final Thesis	New

10.2 Appendix B: Developed functions, algorithms, and code

```python
import cv2
import matplotlib.pyplot as plt
import numpy as np
import rawpy

def vsplit(img):
    """Divides an image into 5 sub-images, sliced vertically.

    Args:
        img (array): Input image.

    Returns:
        array: 5 sub-images, sliced vertically.
    """
    # Divide image into vertical slices
    h, w, channels = img.shape
    vsplit = w//5
    part1 = img[:, 0*vsplit:1*vsplit]
    part2 = img[:, 1*vsplit:2*vsplit]
    part3 = img[:, 2*vsplit:3*vsplit]
    part4 = img[:, 3*vsplit:4*vsplit]
    part5 = img[:, 4*vsplit:5*vsplit]

    return (part1, part2, part3, part4, part5)

def get_top_pxl(img):
    """Returns the location of the top-most non-zero pixel in an image.

    Args:
        img (array): Input image.

    Returns:
        int: The location of the top-most non-zero pixel in an image.
    """
    for r in range(len(img)):
        for pixel in img[r]:
            if pixel[1] >= 20:
                return r

```

```python
def get_bottom_pxl(img):
    """Returns the location of the bottom-most non-zero pixel in an image.

    Args:
        img (array): Input image.

    Returns:
        int: The location of the bottom-most non-zero pixel in an image.
    """
    for r in range(len(img)-1, 0, -1):
        for pixel in img[r]:
            if pixel[1] >= 20:
                return r

```

```python
def get_pixel_vrange(img):
    """Returns the maximum range of non-zero pixels in the vertical direction

    Args:
        img (array): Input image.

    Returns:
        int: Maximum range of non-zero pixels in the vertical direction.
    """
    return get_bottom_pxl(img) - get_top_pxl(img)

```

```python
def get_pixel_hrange(img):
    """Returns the maximum range of non-zero pixels in the horizontal direction.

    Args:
        img (array): Input image.

    Returns:
        int: Maximum range of non-zero pixels in the horizontal direction.
    """
    return get_pixel_vrange(cv2.rotate(img, cv2.ROTATE_90_CLOCKWISE))

```

```python
def hsv_crop(img_hsv, threshold_min = (30, 50, 30), threshold_max = (90, 240, 240)):
    """Segment a given (hsv format) image based on thresholds in the hsv colour space

    Args:
        img_hsv (array): Input image to be segmented (in hsv format).
        threshold_min (tuple, optional): The minimum values for the
            threshold of pixel colours to keep. Defaults to (30, 50, 30).
        threshold_max (tuple, optional): The maximum values for the
            threshold of pixel colours to keep. Defaults to (90, 240, 240).

    Returns:
        array: Input image with all pixels outside of selected hsv range removed.
    """
    # Segment using inRange() function
    hsv_mask = cv2.inRange(img_hsv, threshold_min, threshold_max)
    # Bitwise-AND mask and original image
    img_hsv_masked = cv2.bitwise_and(img_hsv, img_hsv, mask=hsv_mask)
    img_rgb_masked = cv2.cvtColor(img_hsv_masked, cv2.COLOR_HSV2RGB)
    return img_rgb_masked

```

```python
def denoise(img, kernal_size = 10):
    """Applies opening and closing operations using the given square kernal size.
    Removes noise from the image by deleting features that are smaller than the kernal size

    Args:
        img (array): Image to be denoised.
        kernal_size (int, optional): Size of the square kernal used for the opening
            and closing operations. Defaults to 10.

    Returns:
        array: Input image after denoising.
    """
    kernel = np.ones((kernal_size, kernal_size),np.uint8)
    # Erosion followed by dilation. Removes noise.
    img = cv2.morphologyEx(img, cv2.MORPH_OPEN, kernel)
    # Dilation followed by Erosion. Closing small holes inside foreground objects,
    # or small black points on the object.
    img = cv2.morphologyEx(img, cv2.MORPH_CLOSE, kernel)
    return img

```

```python
def manipulate_tags(img, aruco_tag_size_pxls):
    """Corrects perspective of an image so that the camera plane is parallel to,
    and with the same orientation as the surface and corners of an aruco tag.
    Then, crops the image to remove all pixels below the top of the aruco tag.

    Args:
        img (array): An image that includes a 4x4_50 aruco tag.
        aruco_tag_size_pxls (int): The desired size of the aruco tag (in pixels)
            for the output image.

    Returns:
        array: Warped and cropped image.
    """
    arucoDict = cv2.aruco.Dictionary_get(cv2.aruco.DICT_4X4_50)
    arucoParams = cv2.aruco.DetectorParameters_create()
    focal_length, origin, camera_matrix, distCoeffs = init_camera(img)

    # Detect aruco tags
    (corners, ids, rejected) = cv2.aruco.detectMarkers(img, arucoDict, parameters=arucoParams)

    print("ids", ids)
    if ids is not None:
        tags_count = len(ids.flatten())
    else:
        tags_count = 0

    # Determine preference of which tag to warp to (if there is more than one tag detected)
    warp_tag_prefs = [2, 1, 0]
    warp_tag = None
    for id in warp_tag_prefs:
        if id in ids:
            warp_tag = id
    if warp_tag is None: print("Warp tag not found!")

    # Find index for tag to use for image warping
    id_index = None
    for i in range(tags_count):
        if ids[i] == warp_tag:
            id_index = i

    # For the aruco tag, note the current locations of each of the corners
    id = ids[id_index]
    (topLeft, topRight, bottomRight,bottomLeft) = corners[id_index][0]

```

```
    # Compute the desired new corner locations after the image is warped
    pts = np.float32(
        [topLeft,
         [topLeft[0] + aruco_tag_size_pxls, topLeft[1]],  # New top right
         # New bottom right
         [topLeft[0] + aruco_tag_size_pxls, topLeft[1] + aruco_tag_size_pxls],
         [topLeft[0], topLeft[1] + aruco_tag_size_pxls]  # New bottom left
         ])

    # Warp the image to transform corners to new positions
    perspective_transform = cv2.getPerspectiveTransform(corners[id_index][0], pt
    warped_img = warp_image(img, perspective_transform)

    plt.imshow(warped_img)
    plt.show()
```

```
    # Run tag detection again on warped image
    (corners, ids, rejected) = cv2.aruco.detectMarkers(img, arucoDict,
        parameters=arucoParams)
    print("Corners", corners)
    (topLeft, topRight, bottomRight,bottomLeft) = corners[id_index][0]
    print(topRight)
    print(warped_img.shape)
    print(int(warped_img.shape[1]))

    # Crop image below the tag to reduce inaccuracy due to plants overhanging the tray
    cropped_image = warped_img[0:int(topRight[0]+30), 0:int(warped_img.shape[1])]

    return cropped_image

```

```
def get_yield_for_height_at_age(height, age, num_shoots):
    """Uses a known relationship between the height and age of microgreens,
        and the corresponding yield to estimate yield.

    Args:
        height (float): An estimate of the average height of microgreens (in mm)
        age (int): The age of the microgreens (in days since planting)
        num_shoots (int): An estimate of the total number of shoots in the
            population of microgreens

    Returns:
        float: An estimate of the yield of the population of microgreens (in grams)
    """
    yield_per_height_at_age =  (0.00011 * age) + 0.00324
    if age >= 8:
        if age <= 13:
            return yield_per_height_at_age * height * num_shoots

```

```python
def get_pxl_extent(img, num_splits = 5):
    """Splits the image into multiple sections, and returns the average
    (across all sections) range in the x and y directions of all non-zero pixels.

    Args:
        img (array): Input image to get extents for.
        num_splits (int, optional): The number of divisions of the image to average across
            Defaults to 5.

    Returns:
        float, float: Average horizontal and vertical ranges.
    """
    h_pixels_list = []
    splits = vsplit(img)
    for split_img in splits:
        h_pixels_list.append(get_pixel_vrange(split_img))
    print(h_pixels_list)
    h_pixels = sum(h_pixels_list) / len(h_pixels_list) # get average pixel height over ima
    w_pixels = get_pixel_hrange(img)
    return h_pixels, w_pixels

```

```python
def init_camera(image):
    """Initialise the intrinsic properties of a camera.
    It is assumed that there is no lens distortion.
    I.e., the focal length, origin, distortion characteristics, and a transformation matri
        (rotation and translation) describing how a camera maps 3D points in the world to
            2D points in an image.

    Args:
        image (array): An image taken by the camera.

    Returns:
        float: (focal_length):  Estimated focal length of the camera.
        array: (origin):            Vector describing the position of the middle pixel
            in the image.
        array: (camera_matrix):     Transformation matrix describing the rotation and
            translation of the camera.
        array: (distCoeffs):        Matrix describing the distortion characteristics
            of the camera (assuming no distortion).
    """
    # Define camera intrinsic properties (how a camera maps 3D points in the world to
    #   2D points in an image)
    # Matrix can be though of as a rotation matrix concatenated with a translation matrix)
    focal_length = [image.shape[0], image.shape[0]]             # In pixels
    # principal point (the point that all rays converge) in pixels
    origin = np.array([image.shape[0]/2, image.shape[1]/2])
    camera_matrix = np.array(
    [[focal_length[0], 0, origin[0]],
    [0, focal_length[1], origin[1]],
    [0, 0, 1]], dtype="double")
    distCoeffs = np.zeros((4, 1))           # lens distortion of camera (None)
    return focal_length, origin, camera_matrix, distCoeffs

```

```python
def warp_image(img, H):
    """Warps image by transformation matrix (homograph) H.

    Args:
        img (array): Image to warp.
        H (array): Transformation matrix to warp image.

    Returns:
        array: Input image warped by H.
    """
    h,w = img.shape[:2]
    pts = np.float32([[0,0],[0,h],[w,h],[w,0]]).reshape(-1,1,2)
    pts_ = cv2.perspectiveTransform(pts, H)
    xs = []
    ys = []
    for i in range(len(pts_)):
        xs.append(pts_[i][0][0])
        ys.append(pts_[i][0][1])
    xmin, ymin = np.int32(min(xs)), np.int32(min(ys))
    xmax, ymax = np.int32(max(xs)), np.int32(max(ys))

    Ht = np.array([[1,0,-xmin],[0,1,-ymin],[0,0,1]]) # Translate
    result = cv2.warpPerspective(img, Ht.dot(H), ((xmax-xmin), (ymax-ymin)))
    return result

def main():
    # Input parameters
    A_mm2 = 255 * 515      # Total area of the pot
    h_offset = 15            # Distance between the soil and the top of the pot
    age = 10                 # Age in days
    image_path = "./images/"
    image_filename = "image3.jpg"

    # inits
    aruco_tag_size_mm = 49
    aruco_tag_size_pxls = 100
    w_mm = 255               # Pot width (mm) (measured)
    d_mm = 515               # Pot depth (mm) (measured)
    h_tray_mm = 70           # Pot height (mm)
    A_mm2 = w_mm * d_mm      # Total area of the pot

    seed_grams_per_mm2 = 55/(255 * 515)  # Grams per mm^2 (55 grams in a standard pot)
    seed_count_per_g = 690/11            # Seed count/gram (Determined experimentally)
    seed_viability = (420 - 33)/420      # Viable seeds percentage (Determined experimentally)
    num_shoots = A_mm2 * seed_grams_per_mm2 * seed_count_per_g * seed_viability
    num_splits = 5                       # More splits gives more samples for an average
    h_mm_per_pxl = aruco_tag_size_mm/aruco_tag_size_pxls    # Height in mm per height in pixel

```

```
    # Load image (jpg or raw)
    try:    img = rawpy.imread(image_path+image_filename).postprocess()
    except: img = cv2.imread(image_path+image_filename)
    plt.imshow(img)
    plt.show()

    # Warp image to correct for perspective
    warped_img = manipulate_tags(img, aruco_tag_size_pxls)
    plt.imshow(warped_img)
    plt.show()
    #cv2.imshow("Warped", warped_img)
    #cv2.waitKey(0)
    #cv2.destroyAllWindows()

    img = warped_img

    # Convert image to HSV colour space
    img_hsv = cv2.cvtColor(img, cv2.COLOR_RGB2HSV)

```

```
    # Perform image cropping/masking to keep only greens
    img = hsv_crop(img_hsv)
    plt.imshow(img)
    plt.show()

    # Remove noise and fill in holes - all pixels left should be part of plant
    img = denoise(img)
    plt.imshow(img)
    plt.show()

    # Measure average maximum height in pixels
    h_pixels, w_pixels = get_pxl_extent(img, num_splits)

    # Convert pixel height to mm
    h_mm = h_mm_per_pxl * h_pixels
    average_max_height = h_mm + h_offset

    average_height = average_max_height/2

    # Get yield estimate based on height (mm), age (days), and estimated shoot count
    yield_estimate = get_yield_for_height_at_age(average_height, age, num_shoots)

    print("num_shoots", num_shoots)
    print("h_pixels", h_pixels)
    print("w_pixels", w_pixels)
    print("h_mm", h_mm)
    print("yield_estimate", yield_estimate)
    print("-----------------------------")


if __name__ == '__main__':
    main()


```