Automating stress detection in a vertical farming environment using an autoencoder for anomaly detection

Daniël M. van Herwijnen[1], Yossi Capua[2], and Mert İmre[3]

[1]*Author*
[2]*Supervisor, Growy Labs, Amsterdam, the Netherlands*
[3]*Daily supervisor, Growy Data, Amsterdam, the Netherlands, mert@growy.nl*

Abstract

Due to increasing world population and high amount of hunger, better agricultural technology is in demand. Growy is a vertical farming company that is aiming to meet this demand in the future. By scaling-up, methods that could be used before are becoming infeasible, like manually monitoring the health of the crops. Therefore, automated monitoring is required to keep performing well. Chlorosis is easily detectable in microgreens and it is informative about the growth conditions. Thus, automating stress detection could be a perfect way to approach automated monitoring. Stress detection does require the ability to induce stress and capture this stress. These results show that stress can be induced in a vertical farming environment using a tap-water treatment. Capturing stress is also possible, but still requires a lot of improvements to work well in a color sensitive application. Anomaly detection appears to work well for detecting stress, but the experiments should be repeated with more stable and better quality data. These methods provide the foundations for an automated monitoring system for microgreens growing in a vertical farming environment.

Keywords: Anomaly Detection, Computer Vision, Crop Monitoring, Chlorosis, Microgreens, Vertical Farming

1 Layman's summary

Feeding the world is becoming more challenging as the world's population increases and many people are suffering from hunger. Vertical farming could be a way out, because it can improve the amount of food we can produce per square meter. Growy is a very successful company attempting to realize the vertical farming dream by automating the care for plants by robots. However, it is becoming difficult to make sure all plants are growing properly. Currently, the crops are inspected manually, but due to the increase of production, this is becoming very challenging. For this reason, the monitoring process should become automated.

Insufficient nutrients can cause stress in microgreens, which will present by changing of the color of the leaves. Because of this noticeable color change, computer vision could be a good approach to solve the problem. To experiment with this, it is required that you can give plants stress and see the stress in images. We show that stress can be induced, but it is difficult to visualize due to constraints in the image-capturing system. When the system is improved, stress could be detectable using machine learning, anomaly detection, specifically. These results are a first step in automating health monitoring in microgreens growing in a vertical farm.

Utrecht University Student Theses Repository (2023)

Corresponding author
Mert İmre

Edited by
Daniël M. van Herwijnen

2 Introduction

2.1 Need for vertical farming

As global life expectancy has been increasing steadily since 1950 (Ritchie *et al.* 2023), the world's population is predicted to grow to 9.8 billion by the year 2050 (UN 2022). By the year 2050, food demand is said to increase by 60% (Silva 2012). At the same time, 1 in 7 people is suffering from hunger (Silva 2012). These problems demand agricultural technology capable of sustainably and affordably producing food around the globe, like vertical farming due to its space-efficiency. Growy is a young but well performing vertical farming company aiming to bring food cultivation to the next level by increasing sustainability and affordability among others. They grow many different types of crops (i.e. cultivars) like herbs, lettuces and other microgreens. Crops are seeded, mostly automatically, in gutters of 0.10 by 2.40 meters. One hundred of these gutters are laid out next to each other on a growth layer, with multiple growth layers above one another in a growth cell. Each growth layer is serviced by a robot that takes care of daily watering and moves the gutters when and where they are supposed to go. Conditions inside the growth cell can be regulated by adjusting the lights and air conditioning. A vertical lift is being used to transport between different growth cells to make sure the gutters are in the correct conditions during all stages of their growth cycle. After two to three weeks, the gutters exit the growth cell and are harvested, cleaned and prepared to be seeded again and enter a new growth cycle. These farms are cutting edge technology because 1) they decrease space requirement by vertical farming, 2) thereby allow for farming virtually anywhere and 3) automate their farms by robot labor and transportation.

Recently, growy has improved its technology much and are attempting to increase the scale of production and automation in the coming months. As more processes become automated, the attention shifts to more complicated tasks. Now that most basic operations have become automated, the next step is automating the monitoring of farm performance and health. First steps have already been taken, namely by monitoring temperature, CO_2-levels and nutrient contents of the water among others. Another aspect of monitoring is visual inspection, which is an abstract task and is highly dependent on context. During the entire production-cycle from seeding to harvesting, the plants are inspected for aberrant growth. Although manually inspecting the crops is fine at a low scale, it is blocking properly up-scaling because it would require a high amount of manual labor, impeding the goal of affordability. Being able to monitor the health of crops automatically would therefore be a big breakthrough for Growy to reach their goals. In this paper, the focus is on automating stress detection in a vertical farming environment. Stress has been selected specifically because it can indicate whether proper growth conditions are being met. The next section will address stress in more detail to explain why this particular anomaly is selected for this monitoring experiment.

2.2 Stress

Plants can become stressed by nutrient deficiency causing them to display chlorosis, drying or curling among others. Chlorisis is characterized by becoming chlorophyll deficient, causing the leaves to bleach. Chlorophyll is required for photosynthesis, therefore chlorosis can decrease nutritional value of a crop. Chlorosis is caused by insufficiently available nutrients for chlorophyll synthesis (Brown 1961). Among these nutrients, most important are minerals like iron (Abadía *et al.* 2011), and nitrogen (Klotz *et al.* 2016). Additionally, nutrient-uptake can become impaired because of too high or low pH-levels, causing nutrient insufficiency despite nutrient availability (Turner, Arzola, and Nunez 2020). Finally, other problems like improper drainage, root damage or infections can be the cause of chlorosis (Jung *et al.* 2022). Because of a loss of chlorophyll that gives leaves their green color, chlorosis can be identified by bleaching, a process by which green leaves become yellow (figure 1) or white (Jung *et al.* 2022). Humans can therefore classify stress in

Figure 1. Bleached leaves in China Rose radish. Healthy leaves should be looking freshly green like the cotyledons. The secondary leaves are yellow and have a strongly visible venation pattern.

a plant by looking at the leaves. Another important feature of chlorophyll is its ability to perform photosynthesis using blue and red light. As the plant experiences stress, the concentration of chlorophyll decreases in the leaves and thereby, the absorbance of blue and red light decreases. Additionally, near-infrared (NIR) light is strongly reflected by chlorophyll. Therefore, microgreens with stress have a different reflection pattern than healthy microgreens (Jung *et al.* 2022).

For these reasons, the anomaly in microgreens we focus on is stress, because it is 1) easily observable and 2) very informative about the effect of the growth conditions inside a growth cell. Furthermore, in addition to white light, images will also be captured under a NIR and a red light. Images taken under different lighting conditions will be used to train different models to assess which light produces most information-rich images. Together, the experiment could what lighting conditions can be used to automatically detect stress in a vertical farming environment.

2.3 Anomaly Detection

Because chlorosis can be observed easily by a human, it is possible that computer vision can be used to replace humans in this regard. Explicitly teaching a model to classify stress could be a good approach. However, this would require intensive labelling and low potential for re-usability of the model. Since a human detects stress by the outstanding leaves in a gutter, anomaly detection appears to be a good candidate for this task. An anomaly detection model is trained on only normal data, perfect undisturbed growth in this case. After training, it should be able to recognize healthy looking crops and, in contrast, have trouble recognizing microgreens with abnormalities. This method requires no intensive full-image labelling because it is unsupervised learning. Moreover, instead of just being trained to detect stress, anomaly detection can potentially be used to detect more anomalies than just stress, like drying out, damage and other visually presenting anomalies. Therefore, anomaly detection will be used as a model to detect stress.

Autoencoder

Anomaly detection can be implemented like a convolutional autoencoder (Collin and Vleeschouwer 2020). An autoencoder is a two-in-one model, because it consists of an encoder and a decoder. The encoder encodes the input into a smaller dimension, the latent space, and the decoder reconstructs the input from this latent-space-encoding (figure 2. The autoencoder is trained using the mean squared error (MSE)-loss, which calculates the difference between the input and reconstruction. Because the model is only trained on normal looking data, it could have difficulty reconstructing the input containing anomalies because it looks different than the data is was trained on. Therefore, the reconstruction error is expected to be higher for images containing anomalies. This reconstruction error allows for anomaly detection.

Lossy reconstructions

A problem with the MSE is that it leads to blurry outputs (Zhao *et al.* 2015), thereby also not only increasing the error anomalies, but also for normal data. This problem can be circumvented by using skip-connections between the layers of the encoder and decoder. However, additional changes are required, because by introducing the skip-connections, the model can just focus on the skip connection in the last layer, which also contains the input image, nullifying all previous layers. Collin and Vleeschouwer 2020 proposed two changes. First, using an addition-skip rather than a concatenation-skip. A concatenation-skip stacks the channels of two different layers which allows the model to ignore specific parts of the tensor. A mathematical addition prevents the channels from a specific layer to be ignored because the two layers are combined into one. Secondly, Collin and Vleeschouwer 2020 used anomaly simulation to make the input slightly different from the expected output. This forces the model to learn the pattern within the images, because it can occur that part of the image has to be reconstructed entirely. Introducing addition-skip-connections in addition to anomaly simulation should lead to better reconstruction and can therefore lead to a suitable anomaly detecting autoencoder.

Anomaly simulation

Because there are different types of anomalies in the growing cells, we want to experiment with different types of anomaly simulation. Bleaching of the leaves will be simulated by changing the color slightly. Furthermore, seeds of one cultivars can end up in a another's gutter. As a result, part of the leaves of the expected cultivar are blocked by the invading plant. To simulate this, a randomly perturbed elliptical shape with random grayscale values is added to the input image. Finally, two types of noise are also used as anomaly simulation, namely gaussian (Barbu 2013) and salt and pepper noise (Fu *et al.* 2018). These anomaly simulation methods are used to assess whether they can improve the performance of models with skip-connection.

The goal of this project is to investigate whether anomaly detection can be used to detect stress in microgreens in a vertical farming environment. First, this requires the ability to induce and capture stress. To find a good induction method, two different stress induction-approaches are tried; a resource saving short-term treatment and a time and resource intensive long-term approach. For capturing, three different lighting modes, white, red and infrared are combined to see which type of data leads to the best performance. Additionally, the effects of Skip-connections and anomaly simulation are assessed as well, like performed by Collin and Vleeschouwer 2020.

3 Methods

3.1 Stress Induction

Spray treatment

Pieces of the substrate of a gutter mustard frills green were cut off and put in separate boxes. The leaves were sprayed with 20 mL at different concentrations of hydrogen peroxide dissolved in water. The concentrations of hydrogen peroxide ranged from 0.7% to 1.4%. The sprayed solutions were left on the leaves until they evaporated.

Irrigation treatment

Three gutters (control, flush-tap and tap) were sown with wasabi mustard on biostrate. The gutters were germinated for three days in the dark, humid and warm germination cell. During the first two days, the gutters were sprayed with standard germination-solution, which is 0.0070% hydrogen peroxide and 0.0076% nitric acid dissolved in water. At the end of the third day, the gutters were transferred to an automated growth cell where they were irrigated twice daily with 800 grams of AMS2-nutrient. AMS2-nutrient is a 1:1 mix of solution A and solution B, which are described in table 1. After 5 days, the flush-tap and tap gutters were removed from the automatic irrigation and were thereafter twice daily watered manually with 800 mL tap water. To remove most nutrients from the substrate, the flush-tap gutter was first flushed by inclining the gutter and adding a constant stream of tap water for 5 minutes. After 5 days of tap water, i.e. during the weekend, the gutters were watered once per day by the robot using AMS2-nutrient. After two days, the flush-tap gutter was flushed again as described above and both flush-tap and tap gutters were returned to the tap water treatment. In the final two days, the plants did not receive any water. Photo's of all three gutters were collected in the automated growth cell daily with the growth-lights and working-lights off.

Table 1. Nutrient contents of solution A and B. Solution A and B are mixed in a 1:1 ratio to make AMS2-nutrient. The solutions are stored separately to prevent the salts from precipitating.

Nutrients	Solution A (mg/L)	Solution B (mg/L)
Borax	-	5.15
$Ca(NO_3)_2$	710.88	-
$CuSO_4 \cdot 5H_2O$	-	0.56
Iron EDDHA 6%	49.14	-
K_2SO_4	-	48.99
KH_2PO_4	-	306.19
KNO_3	288.24	445.3
$MgSO_4 \cdot 7H_2O$	-	197.18
$MnSO_4 \cdot H_2O$	-	2.16
$Na_2MoO_4 \cdot 2H_2O$	-	0.76
NH_4NO_3	0.12	-
$ZnSO_4 \cdot 5H_2O$	-	2.59

3.2 Data acquisition

Images were collected by a raspberry pi camera module 2 NoIR mounted to the robot. During the night, the growth lights were turned off and the robot performed photoruns. Before a photorun, the gutter was raised 150 mm. A photorun consisted of two image gathering phases, 1) top-view and 2)

Table 2. Number of images per dataset. The data was divided into 6 different groups. The number of images per group can be found in this table.

Dataset	Number of images
Train	75
Test	19
Validation	36
Seedlings	27
Stress	21
Wilted	18
Total	196

side-view. During the top-view, the robot moved from one side of a gutter over the entire length of the gutter in 9 steps. At each of the 9 steps, the robot captured three photo's under different lighting conditions; white, NIR and red. During the second stage, side-view, 8 photo's were captured under only white light condition. For this project, the side-view images were not used. The images were saved locally on the robot together with the corresponding metadata and were uploaded when the robot finished all photoruns of the night. This was done because uploading could sometimes be interrupted and such interruptions should not influence the photorun instructions.

3.3 Preprocessing

Photo's were captured at a resolution of 3280x2464. Due to an alignment issue in the current version of the farm, the top 820 pixels did not contain plants in a gutter (figure 3D and E), but capture the background of a growing layer. In addition, there is an artefact in the images in the form of a strongly lit circle surrounded by an increasing red tint. Therefore, a section of 1300x900 was taken from the middle and the rest of the image was not used.

3.4 Datasets

The images were divided into categories by manual inspection. Images of the first three days were labelled as seedlings since they were very small. Because they were so small, they looked very different from the images after a couple of days. This set is held-out because we can use it as a performance benchmark to show whether the model is able to pick up on very big differences. This set is called the seedlings-set. Images that showed no abnormalities and were not too large were put into the good category. This set was split twice. First the validation-set was made by taking all images from one gutter. The images of the other two gutters were split 80/20% into a train and a test set. Images that appeared to show bleaching or yellow leaves were marked as stressed. Images of drooping plants were marked as wilted due to water deprivation. Healthy leaves that were too large were left out of the train and test set. When it was uncertain whether there was observable stress, the image was left out. During training, images were randomly rotated every epoch to augment the dataset.

3.5 Image input modes

Since there were images under multiple lighting conditions, the models were trained on different combinations of these conditions. All images were initially saved in RGB. RGB was only used for the white light condition. For NIR and red light, the red channel was used. The models were trained on White, White-grayscale, normalized difference vegetation index (NDVI), NIR+Red, White+NDVI, White-gray+NDVI and White+NIR+Red. The + indicates stacking of channels rather than addition.

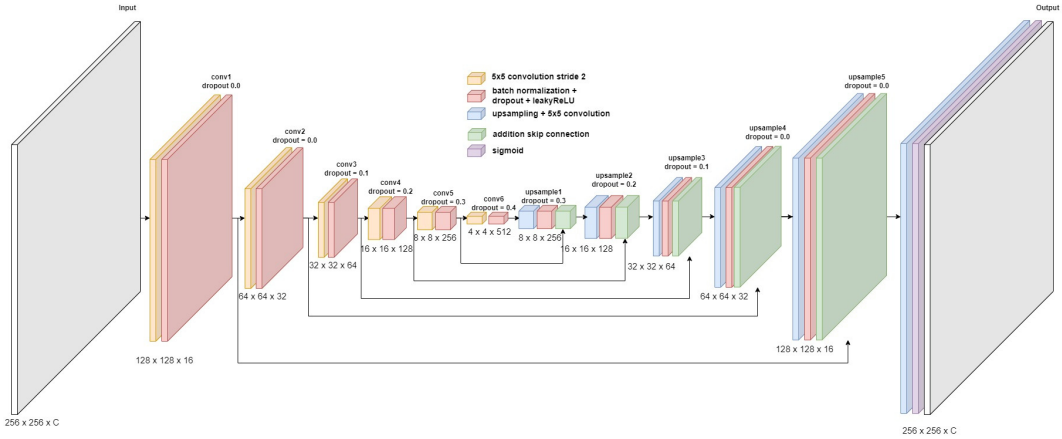


Figure 2. Model architecture of the anomaly detection model. The encoder starts from input and ends at conv6. The decoder starts at conv6 and ends at output. Skip-connections are denoted by black arrows.

The only input mode where the channels where combined is NDVI.

3.6 NDVI transformation

NDVI transformation was applied on an image with red lighting and an image with NIR lighting. For NDVI transformation, only the red channel was used for both. The red channel of the image under red lighting condition was inverted by multiplying by -1 and adding 255, which is the max of an unsigned 8-bit integer. NDVI is defined in equation 1 and has a range of $[-1, 1]$.

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

3.7 Anomaly detection

An autoencoder was trained on images of healthy plants. The autoencoder consists of an encoder-model that is connect to a decoder-model through a bottleneck layer. The encoder decreases dimensions and the decoder increases dimensions (figure 2). The basic building blocks of the encoder consist of a 5x5 strided convolution, batch normalization, dropout and leakyReLU activation. The basic buidling blocks of the decoder consist of upsampling, 5x5 convolution, batch normalization, dropout and LeakyReLU activation. Both encoder and decoder are made up of 6 building blocks, such that the input and output have the same dimensions. For kernel initialization, He-normal was chosen because it is a good fit when using LeakyReLU (He *et al.* 2015). The network was trained using mean squared error between the input and output. Therefore, the aim is that the decoder is trained to recreate the input based on the output of the encoder.

3.8 Model performance

To assess the model performance, the residual sum of squares (RSS) was calculated between input and prediction for each image in each dataset. High residual corresponds to high error and low residual to low error. The RSS is defined in equation 2.

$$\text{Residual sum of squares} = \sum_{i=1}^{n} (P_i^{input} - P_i^{prediction})^2 \tag{2}$$

where n is the number of pixels P in the input.

To compare performance of models with a different number of input channels, the average was taken over the channel dimension. Using the residual, performance of the model was characterized

using receiver-operator-characteristic (ROC)-curves and precision-recall (PR)-curves.

4 Results

4.1 Stress Induction

Spray treatment

At concentrations below 0.35%, no effect of the spray treatment could be observed. Leaves sprayed with higher concentrations displayed some minor blistering (figure 3C). This blistering-effect was very localized to were the peroxide drops had landed. The observed effect did not resemble bleaching. In addition to blistering, the plants also presented with signs of mildew. This was visible by strong discoloration in the lower leaves.

Irrigation treatment

After a few days, a difference between the gutters was observable by their size. The gutter receiving normal treatment had grown fuller than the two gutters on the tap-water treatment. After two weeks, the first signs bleaching leaves were observable when taking the gutters out of the cells. During the final week, chlorosis was clearly visible under lighting conditions outside of the growth cell by bleaching and the venation pattern (figure 3F).

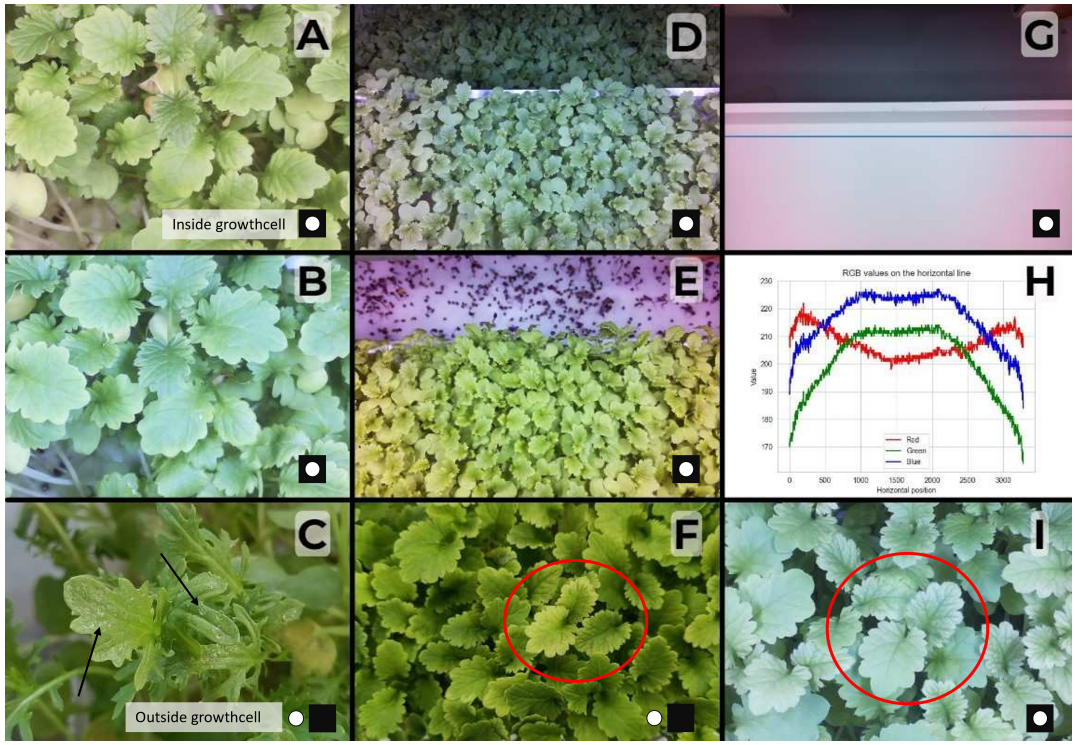


Figure 3. Wasabi Mustard captured inside and outside the growth cell at different days and different positions. Images inside growth cell were taken by the robot camera and images outside the growth cell with a cellphone camera. **A)** Inside growth cell day 18 cropped to center. **B)** Same as A on day 19. **C)** Blistering after spraying with high concentration of hydrogen-peroxide captured outside growth cell. **D)** Inside growth cell day 12. **E)** Same day, position and light as D, but different gutter. **F)** Chlorosis captured outside growth cell on day 19. **G)** Empty gutter captured under white light showing a circular artefact in the center. The blue line in the center shows which pixel-values are plotted in H. **H)** Plot of the RGB values of the pixels on the blue line from G. **I)** Captured roughly the same position as F inside growth cell with white light on day 17.

4.2 Image Capturing

Visible stress was difficult to capture

Images inside the growth cell were collected daily. This allowed the growth to be tracked over time. The main goal was to capture stress using this system. Images were captured inside and outside the growth cell during the days that stress was observable. Figure 3F shows stress captured outside the growth cell with a smartphone camera. Bleaching was clearly visible as well as the venation pattern. An image of stress captured inside the growth cell can be seen in figure 3I. A venation pattern appeared to be distinguishable, but there was no such clear distinction in color as in figure 3F. Additionally, there was an overall color difference between the images captured inside and outside the growth cell. The images inside the growth cell appeared much more blue/white than the images outside the growth cell.

Images varied over time and position in the cell

The images captured inside the growth cell were compared among each other. Figure 3A and B show an image of the same position of the same gutter, captured at day 18 and 19, respectively. The leaves displayed very different colors, 3A more green/red and 3B more blue/white. Additionally, images from the same day taken of different gutters but equal X-coordinates were compared as well. The gutter shown in figure 3D was in between the two wasabi mustard gutters from the irrigation treatment experiment. The gutter shown in figure 3E is visible in the background of 3D. The other gutter shown in figure 3E was next to the previous gutter and next to a recently seeded gutter that did not belong to the irrigation treatment experiment. There is a clear difference between the colors of figure 3D and E, where D is much more blue/white and E is more green/red. This difference was the same over the entire length of the gutters. Finally, multiple images were taken of the same gutter and the same x-position on the same day (figure 4). Differences in brightness and color are clearly visible between these images. These difference were not visible when the growth lights are on.

To verify stability of the camera system, multiple images of the same position were captured in succession under white light. The red, green, blue and brightness values were plotted over time (figure 4K and L). The values are stable over time for the images captured with the growth lights on. In contrast, without growth light, the values vary strongly between the first three images. This is clearly observable between the first and last images (figure 4F and J).

In addition to the white images, color difference are also visible in the images taken under NIR and red light. The red images can be grouped into images that look a more deep red and images that have a brighter red (figure 1A and B). For the NIR images, the images can be grouped into mostly pink and more blue/purple (figure 1D and E). The difference between the NIR images is less apparent than the red images (supplementary figure 1).

Artefact

In addition to color differences between images, a circular artefact is visible in all images. This artefact is most apparent on images of an empty gutter (figure 3G). The center of the circle shows strong white reflection, whereas the edges of the images are more red. The RGB-values along the horizontal blue line are plotted in figure 3H. In the center, the RGB-values are stable, whereas they change towards the edges. The green and blue values start to decrease at the same horizontal position and appear to decrease the same amount; the difference between green and blue values appears to be constant. In contrast, the red values increase when moving away from the center and finally drop as well. The artefact is also clearly present in the images taken by NIR and red light (supplemental figure 1). For infrared (F), there is a hotspot of light in the center of the image, whereas the image taken under red light (C) shows the same red increase further from the center.

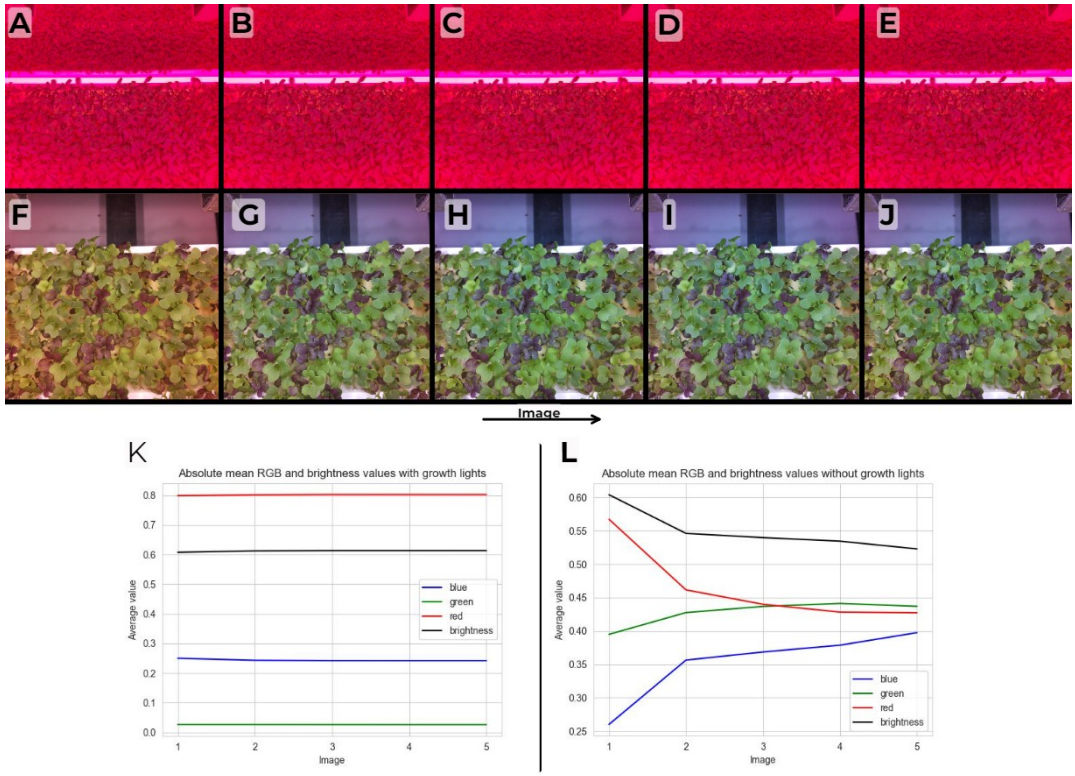


Figure 4. Inconsistent image sensor startup. Multiple images were captured in short succession. **A-E** Chronologically ordered images captured in short succession with the growth lights on. **F-J** Chronologically ordered images captured in short succession with the growth light off. **K** Average RGB- and brightness values of images A-E. **L** Average RGB- and brightness values of images F-J.

4.3 Anomaly Detection

Models with skip-connection were trained on the train- test-, validation and seedlings-dataset to assess how well it was able to learn what normal wasabi mustard looks like. The models were different in the type of colormode they used. For all models, the train- and test-set have a similar low residual and the seedlings-set is much higher for all models. Most models have a higher residual for the validation set, but still much lower than the seedlings-set. The model trained on white-grayscale shows an equal residual between train, test and validation.

The ROC- and PR-curves show perfect performance for the models trained on white-gray and white-gray-NDVI. The model trained on NDVI performs worse than the one trained on red and infrared. The model trained on white-rgb shows better performance than the model with other image-types in addition to white-rgb.

In an attempt to further improve the difference between train- and test-data and validation-data, four models with skip-connections were trained with different types of anomaly simulation. Among the models with anomaly simulation, salt and pepper noise was the best model. However, none of the models had significantly better performance than the model without anomaly simulation.

Skip models were trained with data from different lighting conditions. The model trained on color images taken with white light outperformed almost all other models significantly. The model trained on white grayscale + NDVI images performed worse, but not significantly. The two worst performing models were the model trained on NDVI and the model trained on NIR+Red.

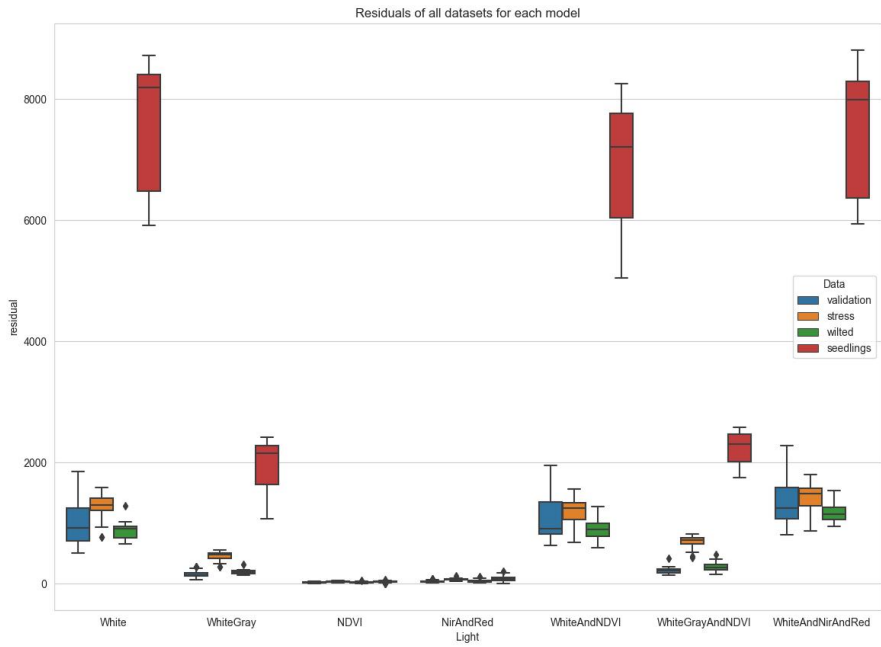


Figure 5. Residuals for each image in each dataset. The residuals are separated by their respective models. Residuals of the seedlings set is very high for each model. Residuals of the stress dataset are lower than seedlings, but still high for some models.

5 Discussion

The goal of this project is to assess whether stress in microgreens can be detected using computer vision. This requires the ability to induce stress in microgreens and then capture the stress using cameras. We were able to induce stress using the tap water treatment but not by spray treatment. Capturing the stress in photo's appears to be possible but still needs improvements. The anomaly detection model seems to be able to reconstruct images of healthy microgreens and has a high RSS for images displaying anomalous growth in microgreens.

5.1 Stress induction

To train a model capable of detecting stress in microgreens, images of stressed microgreens are required. We tried inducing stress long- and short-term. A hydrogen-peroxide-spray treatment was used for short-term stress induction. The treated microgreens developed blisters on the leaves, but this did not resemble chlorosis. Therefore, this approach cannot be used to induce the stress we are looking for. However, only hydrogen-peroxide has been used to attempt to induce stress. Possibly, using a different chemical could induce stress through a spray treatment. For example, a strong acid might be a candidate to try next. The benefit of spray treatment is that it could be much faster and less resource intensive than irrigation treatment. Therefore, I suggest conducting more experiments with different chemicals. In contrast to the spray treatment, the long-term tap water treatment was able to induce stress in wasabi mustard in two weeks. This approach has only been tried on wasabi mustard and can therefore not be guaranteed to also induce stress in other cultivars. For this reason, this approach should be repeated for other cultivars. In addition, tap-water can still contain some nutrients for microgreens and might therefore not be the most optimal stress-inducer. I propose that the same treatment is also performed with reverse-osmosis-water

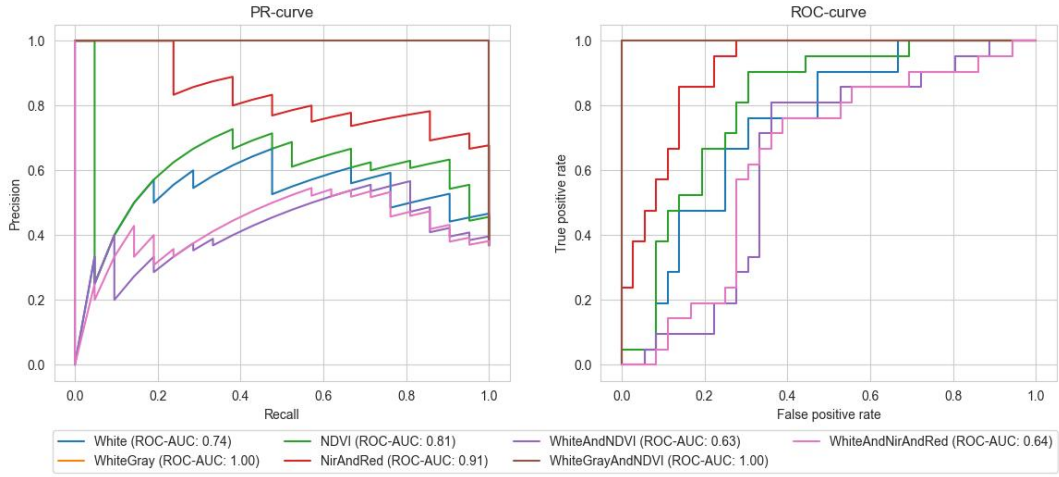


Figure 6. Performance of models using residual For each model, the residuals of the validation and stress data is used. When increasing the threshold, the performance changes. Perfect PR-curve starts at (0,1), moves to (1,1) and then (1,0). Perfect ROC-curve starts at (0,0), moves to (0,1) and then (1,1).

instead of tap water because more nutrients can be removed using reverse-osmosis. Therefore, reserve-osmosis-water could be a more potent stress-inducer than tap water.

5.2 Image capturing

Images of healthy and stressed microgreens were captured under white, red and infrared light. It was more difficult to identify stress in images under these lighting conditions than it was outside the growth cells. The first cause is that images appear very blue/white. This makes it difficult to detect color changes. Importantly, the images have not been corrected by color calibration or other post processing. Doing this could possibly improve the usability of the images. For example, using a reference card with known color-values, the camera can be calibrated to produce more true colors. This could drastically improve image quality which could be beneficial for color-sensitive applications, such as stress detection.

Canopy height

However, the second cause of aforementioned difficulty is image inconsistency, which also complicates calibrating the camera sensor. The results presented here show that images taken on different days can differ drastically (figure 3A and B). Color differences could in part be explained by the proximity of the canopy to the camera. Since the gutter are lifted to a fixed height before capturing, the canopy is closer to the camera the taller it grows. Following the law of squares, more light will hit the canopy increasing the brightness produced in the image. Therefore, lighting is possibly inconsistent because of canopy height.

Background

Nevertheless, our data show that height cannot be the only cause of image inconsistency. Images gathered at equivalent canopy height also show differences. It appeared that these differences could be caused by the amount of white in the background. The images belonging to the gutter shown in figure 3E all had a white background and were all more green than the images belonging to the gutter from figure 3D. The different appearance could be explained by automatic white balancing performed by the image sensor. This should be verified by disabling that feature and properly calibrating it, for example using a reference card. Fixing the white balance could further improve the quality of the images gathered inside the growth cell.

Sensor saturation

Furthermore, the saturation of the camera sensor also appears to have been causing inconsistency. Figure 4L clearly shows that the images gathered over time are not consistent. The exact cause of this effect is not known to me, but could be caused by sensor saturation, or a different feature of the camera sensor changing the image. Anyway, this also affects the consistency of the images and therefore needs to be addressed for color sensitive applications.

Other causes

So far, mainly problems on the camera side have been discussed. However, other issues could possibly affect image-gathering inside the growth cell. The LEDs are controlled by changing their voltages. Previously, it has been shown that the LEDs can have power issues (Dawe 2022). From the data presented in this paper it cannot be excluded that power issues are causing the observed differences. The possibility of this problem needs to be examined in the future. Furthermore, figure 4K shows that images captured while the growth light were much more consistent then without the growht light. In addition, the optimization of the number of LEDs was performed by eyeballing. Therefore, it could be possible that the number and type of LEDs needs more optimization to get images of proper quality.

A problem with optimizing all these different parameters is that the experiment runs inside a production environment. It can be convoluted to work perform an experiment that would seem easy. This is because the camera module is installed on the robot that is tasked with the day-to-day care for all crops. Experiments can be interrupted by tasks the robot has to perform for other departments in the company. Additionally, the robot is much more elaborate than just a camera system. This can raise issues that are completely unrelated to the camera but still hinder experiments with it. Finally, it was not allowed to move crops to the production cell if they were grown elsewhere. Certain types of anomalies can be found during harvesting and it would be beneficial to capture images of these anomalies. However, since the camera system is inside the growth cell, this is not possible. Therefore, it would be much more ideal to separate the basic optimization stage from the production-ready product. If a more accessible version of the camera system would be available, experimenting with lights, calibration, camera settings and gathering relevant data would be much easier and could be achieved much faster.

Camera type

The current camera captures images in color and applies filters in the process. These RGB-cameras do have the benefit that they produce images that a human can understand very easily because they are very similar to our visual system. Other researchers use monochrome cameras instead because they do not have to apply filters (Steven 2023). Monochrome capturing also opens the possibility to multi-spectral imaging which can be used to assess the chemical-contents of microgreens (Dieleman *et al.* 2018). Therefore, I propose to experiment with monochrome cameras instead.

Regardless of the type of camera used, it appears that the current positioning of the LEDs leads to a circular artefact on the images. This artefact makes a large part of the image difficult to use for analysis. Spreading out the LEDs over the width of the field of view could potentially circumvent this issue. However, the artefact might also be caused by the fish-eye effect. For color- and light-sensitive application, the artefact needs to be removed.

5.3 Anomaly Detection

Anomaly detection aims to reproduce the input by autoencoding. During inference using a perfect model, a low RSS is expected for normal data and a high RSS for data with anomalies. It has been shown that MSE leads to blurry images and that this could be circumvented using an addition-skip connection (Collin and Vleeschouwer 2020). The data suggests that this also holds for our

data. Introducing a skip-connection into the model leads to much lower RSS in general, but the decrease is most noteworthy for the train- and test-data. Therefore, the difference between normal and abnormal data becomes larger by using a model with a skip-connection, allowing for better anomaly detection.

The train- and test-data give similar RSS for all models. The seedling-data has a very high RSS, which indicates that the input in not simply transferred through the skip-connection. Additionally, the validation data has a much lower RSS than the seedling data. Furthermore, the stress data leads to a higher residual than the validation data. Together, these findings appear to indicate that the model is able to learn what healthy wasabi mustard looks like and that the model struggles to reconstruct the image when stress is present. Therefore, it appears that this model is suitable for anomaly detection in microgreens.

Models trained on grayscale images taken by white light seem to lead to the best performance, which is in line with the work of Collin and Vleeschouwer 2020. Single channel input works well for anomaly detection. For this reason, training our model on NDVI is expected to perform well. Strikingly, stacking the red and infrared image into a 2-channel image appears to lead to a better model than when training on NDVI. Models trained on white-rgb did not improve when adding other image types, like NDVI or red and infrared. From this, we conclude that grayscale images captured under white light are best suited for training anomaly detection. However, this could possibly change when the image capturing system is improved using aforementioned changes.

One important difference between the seven different input-modes is the number of channels in those inputs. White-grayscale has only one channel, whereas White-rgb+NIR+red has five channels. To compare these models, the average residual was calculated over the channels. This could have caused the influence of an important channel to have been lost. However, since the models trained on 1) NDVI and 2) Red+NIR did not perform well, this appears to be unlikely. Therefore, it is not likely that taking the average over the channels has caused a bias.

Since stress was not as obviously distinguishable in the images gathered in the growth cell, there is the possibility that the images selected for containing stressed leaves did not actually contain stress. Therefore, the model would not have been able to learn to detect stress since it might not have been present in the data. In the future, lighting conditions that properly light the microgreens are required to be able to detect chlorosis like is possible outside the growth cell.

Experiments were also conducted with models trained using anomaly simulation. None of the anomaly simulations improved the performance significantly when trained on RGB or grayscale images taken under white light.

The models were trained on a dataset with a low number of images. This was mitigated by augmentation through random rotation. However, more data would be beneficial to make the model much stronger. This can be achieved on one hand by gathering images of many more gutters. Additionally, if the circular artefact can be removed, larger parts of the image become available for training. Alternatively, expanding augmentation with random brightness or contrast could possibly increase the amount of usable data without taking more images. It is possible that the pixels tainted by the artefact are usable when these changes are made to augmentation. However, on the currently used limited dataset, the model already generalized rather well on the test- and validation-set. So, the model could become stronger and more robust, but the results presented in this paper are still good as well.

All cultivars

All models were only trained on images of wasabi mustard. This cultivar was chosen because it can be grown easily and stress can be seen clearly. However, Growy has many more crops than just wasabi mustard. To detect anomalous growth in other cultivars, data needs to be collected for each and a model has to be trained on this data. The easiest step is to train and deploy a copy of

the model described in this paper for each cultivar. Because gutters are equipped with an IR-chip, the type of crop can be read-out by the robots and the corresponding model retrieved. Another solution is to train a single model that can learn to distinguish cultivars and then perform anomaly detection. A slight variation on the auto-encoder presented here could potentially be used for this application: a variational auto-encoder (Hsu, Zhang, and Glass 2017). I propose the name cultivariational auto-encoder. In theory, the model's latent space could also be used to perform anomaly detection. If all cultivars can be fit in this latent space with as little overlap a possible, the auto-encoder could be able to learn the growth pattern of multiple cultivars at the same time. Distributing the cultivars on the latent space could possibly be achieved by the reparameterization trick that characterizes the variational auto-encoder. Whether additional changes to the current model architecture are required is difficult to predict. The first step should be to get very reliable data-collection and retry stress detection using anomaly detection, only thereafter implement reparameterization and train the model on more than one cultivar.

6 Conclusion

Automatic monitoring in vertical farming is required for quality assurance when scaling up production. Cameras can be used to gather data from the crops. However, for advanced tasks such as color-sensitive stress detection, the image quality needs to be improved much more. This could be achieved by color calibration, LED and camera sensor optimization and understanding the lighting characteristics inside a growth cell. Despite the suboptimal image quality, an autoencoder with skip connections appears to be able to learn what healthy wasabi mustard looks like and can be used for anomaly detection. Anomaly detection appears to be a promising technique for monitoring the desired growth of microgreens. When the image capturing system is improved, this experiment should be repeated and expanded upon by implementing a variational autoencoder instead. Hopefully, this technique can then be used to monitor the entire growing process inside a vertical farm to ensure, healthy, tasty and affordable food around the globe.

References

Abadía, J., S. Vázquez, R. Rellán-Álvarez, H. El-Jendoubi, A. Abadía, A. Álvarez-Fernández, and A. F. López-Millán. 2011. "Towards a knowledge-based correction of iron chlorosis." *Plant Physiology and Biochemistry* 49 (5): 471–482. ISSN: 0981-9428. https://doi.org/10.1016/J.PLAPHY.2011.01.026.

Barbu, T. 2013. "Variational image denoising approach with diffusion porous media flow." *Abstract and Applied Analysis* 2013. ISSN: 10853375. https://doi.org/10.1155/2013/856876.

Brown, J. C. 1961. "Iron Chlorosis in Plants." *Advances in Agronomy* 13 (C): 329–369. ISSN: 0065-2113. https://doi.org/10.1016/S0065-2113(08)60963-3.

Collin, A. S., and C. de Vleeschouwer. 2020. "Improved anomaly detection by training an autoencoder with skip connections on images corrupted with Stain-shaped noise." *Proceedings - International Conference on Pattern Recognition,* 7915–7922. ISSN: 10514651. https://doi.org/10.1109/ICPR48806.2021.9412842.

Dawe, A. R. 2022. *Vertical Farming Sensor Integration and Evaluation (Unpublished Manuscript).* Technische Universiteit Delft, November.

Dieleman, A., G. Polder, E. Meinen, J. V. Arkel, and K. Weerheim. 2018. "Klimaat sturen op de inhoud van het blad Rapport WPR-770," https://doi.org/10.18174/461263. www.wur.nl/plant-research..

Fu, B., X.-Y. Zhao, Y.-G. Ren, X.-M. Li, and X.-H. Wang. 2018. "A salt and pepper noise image denoising method based on the generative classification" (July). https://arxiv.org/abs/1807.05478v1.

He, K., X. Zhang, S. Ren, and J. Sun. 2015. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification."

Hsu, W. N., Y. Zhang, and J. Glass. 2017. "Unsupervised domain adaptation for robust speech recognition via variational autoencoder-based data augmentation." *2017 IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2017 - Proceedings* 2018-January (July): 16–23. https://doi.org/10.1109/ASRU.2017.8268911.

Jung, J., J. H. Baek, Y. Lee, S. E. Jeong, and C. O. Jeon. 2022. "The Self-Bleaching Process of Microcystis aeruginosa is Delayed by a Symbiotic Bacterium Pseudomonas sp. MAE1-K and Promoted by Methionine Deficiency." *Microbiology Spectrum* 10 (4). ISSN: 21650497. https://doi.org/10.1128/SPECTRUM.01814-22/SUPPL_FILE/SPECTRUM.01814-22-S0002.XLSX. https://journals.asm.org/doi/10.1128/spectrum.01814-22.

Klotz, A., J. Georg, L. Bučinská, S. Watanabe, V. Reimann, W. Januszewski, R. Sobotka, D. Jendrossek, W. R. Hess, and K. Forchhammer. 2016. "Awakening of a Dormant Cyanobacterium from Nitrogen Chlorosis Reveals a Genetically Determined Program." *Current Biology* 26 (21): 2862–2872. ISSN: 09609822. https://doi.org/10.1016/j.cub.2016.08.054.

Ritchie, H., L. Rodés-Guirao, E. Mathieu, M. Gerber, E. Ortiz-Ospina, J. Hasell, and M. Roser. 2023. "Population Growth." *Our World in Data* (July). https://ourworldindata.org/population-growth.

Silva, J. G. D. 2012. "Feeding the World Sustainably | United Nations" (June). https://www.un.org/en/chronicle/article/feeding-world-sustainably.

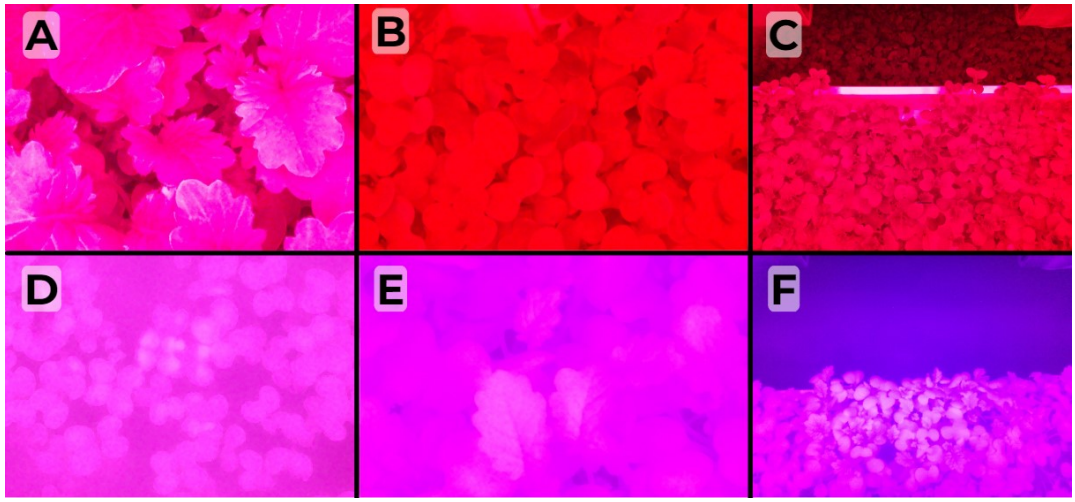
Steven, v. d. B. 2023. *Lectoraat Fotonica.* https://www.dehaagsehogeschool.nl/onderzoek/lectoraten/fotonica.

Turner, A. J., C. I. Arzola, and G. H. Nunez. 2020. "High pH Stress Affects Root Morphology and Nutritional Status of Hydroponically Grown Rhododendron (Rhododendron spp.)" *Plants* 9 (8): 1–12. ISSN: 22237747. https://doi.org/10.3390/PLANTS9081019. /pmc/articles/PMC7465443/%20/pmc/articles/PMC7465443/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7465443/.

UN. 2022. *World Population Prospects - Population Division - United Nations.* https://population.un.org/wpp/Download/Standard/MostUsed/.

Zhao, H., O. Gallo, I. Frosio, and J. Kautz. 2015. “Loss Functions for Neural Networks for Image Processing” (November). https://arxiv.org/abs/1511.08861v3.

A Supplementary Material



Supplementary Figure 1. NIR and red images A-C are images captured under red light. Images D-F are captured under NIR.