agriculture

MDPI

Article

Machine-Learning Approach to Non-Destructive Biomass and Relative Growth Rate Estimation in Aeroponic Cultivation

Oskar Åström [1], Henrik Hedlund [2] and Alexandros Sopasakis [1,*]

[1] Department of Mathematics, Faculty of Science, Lund University, 221 00 Lund, Sweden
[2] Alovivum AB, Göingegatan 6, 222 41 Lund, Sweden
* Correspondence: alexandros.sopasakis@math.lth.se

Abstract: We train and compare the performance of two machine learning methods, a multi-variate regression network and a ResNet-50-based neural network, to learn and forecast plant biomass as well as the relative growth rate from a short sequence of temporal images from plants in aeroponic cultivation. The training dataset consists of images of 57 plants taken from two different angles every hour during a 5-day period. The results show that images taken from a top-down perspective produce better results for the multi-variate regression network, while images taken from the side are better for the ResNet-50 neural network. In addition, using images from both cameras improves the biomass estimates from the ResNet-50 network, but not those from the multi-variatemultivariate regression. However, all relative growth rate estimates were improved by using images from both cameras. We found that the best biomass estimates are produced from the multi-variate regression model trained on top camera images using a moving average filter resulting in a root mean square error of 0.0466 g. The best relative growth rate estimates were produced from the ResNet-50 network training on images from both cameras resulting in a root mean square error of 0.1767 g/(g·day).

Keywords: machine learning; aeroponics; hydroculture; neural network; regression; biomass; fresh weight; relative growth rate; image analysis

check for updates

Citation: Åström, O.; Hedlund, H.; Sopasakis, A. Machine-Learning Approach to Non-Destructive Biomass and Relative Growth Rate Estimation in Aeroponic Cultivation. *Agriculture* **2023**, *13*, 801. https://doi.org/10.3390/agriculture13040801

Academic Editors: Sebastian Kujawa and Gniewko Niedbała

Received: 17 February 2023
Revised: 16 March 2023
Accepted: 21 March 2023
Published: 30 March 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In conventional farming, the soil is used to transfer nutrients to the plant. Although this is possibly the cheapest approach, it is not necessarily the most effective for plant growth or even the most environmental. Large amounts of water, for instance can be lost into the surroundings due to evaporation and dispersion into groundwater [1]. Furthermore, fertilizers used in field farming are also transported away from their intended targets to the environment while leading to nutrient waste and water eutrophication [2]. In addition, farming on fields leaves the plants exposed to a number of harmful factors, such as droughts, soil-borne diseases, bad weather, pest attacks and floods [1,3].

An alternative to soil-based farming is hydroculture, where the nutrients are carried to the plants using water. In particular, hydroponic farming has seen a large increase in commercial use in the last few years [4] with companies, such as Nordic Harvest, Swegreen and Gronska, being some recent Nordic examples.

Aeroponic cultivation is another soil-less farming method that also uses water as the nutrient carrier. Unlike regular hydroculture, however, which uses flowing water, aeroponics uses an aerosol to deliver the nutrients. The first direct benefit of this approach is reduced water usage, which is of great interest for several climate affected regions finding themselves under drought conditions. As a comparison, the water usage of 1 kg of tomatoes is 200–400 L in conventional soil farming, 70 L in hydroponics and 20 L in aeroponics [5].

Aeroponics, however, has an advantage over hydroponics in terms of plant growth rates. A limiting factor of hydroponic farming is that the oxygen content in the water must

be a maximum of 8 ppm. There are no such constraints, however, in aeroponics and it has been shown [6] that the aeration of the plant roots is much better leading to a higher growth rate.

The current emergence of hydrocultural practices, such as aeroponics, is one of many possible paths toward securing sustainable food production in the future. The ability to have controlled growth conditions, which such practices entail, makes it possible to optimize the cost-to-yield in any number of different metrics of importance: water, nutrition, land-use, emissions, or finances. However, in order to make these optimizations, it is necessary to determine the effects that parameters, such as temperature, pH, lighting, space, CO_2, and nutrient concentration, have on the development of plants. These experiments often require destructive measurements of the plant's characteristics by harvesting them. This prevents the plant from growing further and only results in one data point per plant. This method is not only labor intensive but results in poor estimates on an individual level and can only contribute to an estimate of the population as a whole. In order to estimate the effects on the scale of individual plants, continuous and non-destructive biomass estimates are the preferable choice.

The goal of this work is to evaluate two different image-based machine-learning methods for estimating plant growth in aeroponic farming. These two methods are: multivariate regression (MVR) and neural networks using a pre-trained ResNet-50 (R-50) network. The aim is for these methods to estimate the biomass (g) of a given plant as well as the Relative Growth Rate (RGR) measured in $g/(g \cdot day)$. With a sufficiently accurate estimate of RGR, a virtual sensor could be developed to measure the RGR continuously. This would allow for real-time optimization of plant growth, which could decrease the amount of unnecessary resources spent, including water, nutrients, lighting, electricity and labor.

A high RGR is desirable both for industry and research [7] and non-invasive methods that can achieve that result are desirable. It is well-known, however that [8] RGR varies between plant individuals. As a result, estimating the RGR requires multiple time data points per plant. Measurements based on removing parts of the plant can impact or even stop growth. Destructive methods for measuring RGR are therefore not useful for practical industrial applications and may be the reason behind the lack of studies in this area. Non-destructive machine learning approaches such as those we propose here are therefore needed.

We begin in Section 2 with the state of the art in this field. Both of the methods used, the multi-variate regression network and the ResNet-50-based neural network, as well as the image collection practices used are presented in Section 3. We present our results in Section 4. We discuss these results in Section 5 and end with our conclusions in Section 6.

2. State of the Art and Challenges

While the modern incarnations of aeroponics have existed since the mid-1900s [9], the use of aeroponic platforms in conjunction with machine learning (ML) is still in its infancy [10]. As recently as 2022, in the field of deep learning, there were only "14 publications on hydroponic agriculture, one in aquaponics and none on aeroponics for soil-less growing media" as reported in [10]. There are for instance, some studies [11,12] that employed image-to-biomass estimation in hydroponic cultivation. In that respect, it is now possible to find public datasets for training machine-learning models although still from a single (top-down) viewpoint [13]. There exist studies in aeroponics that apply machine-learning methods in order to predict yield from manually measured features, such as the number of leaves and stem diameter [6]. However, there do not exist any studies investigating the potential of image analysis as a tool for biomass prediction in aeroponic cultivation [10,14].

In studies analyzing plant growth using machine learning, it is common to use a classification output where the plant growth is lumped into a number of stages of progression [7,15,16], corresponding to different visual and biological processes in the plants. This makes the learning easier since the targets are limited to discrete values, and usually one is

only interested in a rough estimate of the plant's state of growth. However, for the purpose of estimating the relative growth rate of biomass, we need to be as precise as possible, and therefore a numeric value of the biomass is necessary.

The success or failure of ML approaches depends heavily on the available data. For some applications, collecting the needed data may result in gradually destroying the source of the data at the same time. This is the case in several cell-staining applications where, in order to understand the structures in the cell, one must apply chemicals that illuminate these structures and also destroy them at the same time. These issues also occur while collecting the data needed to measure biomass in plant cultivation studies. Classical approaches typically require data to be collected through weighing by removing leaves. This is a highly manual task, which naturally also results in a low frequency of data points.

In these cases, reported collection frequencies range from hree to four times a week [11] to once at the end of the growth period [12]. This results in a low number of data points and provides less information about growth patterns. In order to achieve a higher time resolution, we exploit a known fact in plant physiology. It has been shown [17] that, in controlled environments where nutrients are provided with free access, the relative growth rate (RGR) is constant. This enables the interpolation of biomass between data points, reducing the need for frequent data gathering and increasing the accuracy of biomass estimation methods.

Another challenge in most approaches dealing with estimating plant growth from images is occlusion [11,12]. Plants tend to obscure other plants as they grow, making individual plant growth more difficult to estimate exactly at the moment it is needed the most. Occlusion, which is thought of as an image data-collection problem, indirectly implies that plants may, in fact, be affecting the growth rate of other plants. For this reason, we consider and analyze the effects of having multiple viewpoints as input.

3. Data Processing and Methods

The main goal of this paper is to compare machine-learning methods for estimating the biomass of a plant given a set of images. The first method that we consider is Multi-Variate Regression (MVR), which employs linear regression on a set of manually curated non-structural features. This is similar to methods employed by Jung, Dae-Hyun, et al. in [12] for estimating biomass in hydroponically grown lettuce. The second approach we consider is based on a convolutional neural network (CNN) architecture, where we furthermore employ transfer learning with a pre-trained ResNet-50 neural network. We refer to this method as R-50. This method is similar to methods employed by N. Buxbaum et al. for hydroponically grown lettuce in [11].

Furthermore, two different models are generated and trained for each method, one for the top and one for the angled camera. The average of the estimates produced by these models is then used to construct a third estimate, called the Dual View. The resulting biomass estimates generated by these three views on the test set are then compared in three tasks:

- Task: Single Image. Biomass estimation from a single time point.
- Task: Moving Average. Biomass estimation from three consecutive time points.
- Task: RGR. Relative growth rate estimation from three randomly sampled time points.

The root mean square error (RMSE) of the results on these tasks is used to compare the different methods.

3.1. Experiment Setup

The setup used for plant growth and development consists of four growth beds and one reservoir. The reservoir contains nutrient-rich water, along with sensors for temperature, pH, and electrical conductivity and a heater to keep the water at a preferred set temperature. The water from the reservoir is then pumped to the four growth beds.

Each growth bed consists of a container with a removable lid. Inside the container, there are two sonicators, which contain membranes that vibrate at an ultrasonic frequency

that agitate the water into an aerosol, in the form of dense fog. The lid of the container has 24 holes with small baskets (plant holders), in which the plants are placed such that the roots hang down and are immersed in the aerosol, continuously replenishing the supply of water and nutrients. The bottom of the container has a drainage pipe that returns the water back to the reservoir, as shown in Figure 1. The plants were grown over a 5-day period, with 57 plants completing the full cycle.

The camera rig consists of 8 cameras (2 per bed) of the model PlexGear WC-800 (manufactured by PlexGear, Malmö, Skane, Sweden) set up as shown in Figure 1. Four of the cameras capture images from a top view, placed on the long edge of each bed. These cameras are referred to as the top cameras. The other four cameras are placed further down and over the adjacent bed so that they capture images from a lower angle. These cameras are referred to as the angled cameras.

The cameras are connected to a computer next to the rig, which runs a script to capture images with each camera at 1 h intervals. These images are then stored locally and sent directly to a cloud-sharing service so that the camera status and plant development can be monitored remotely.

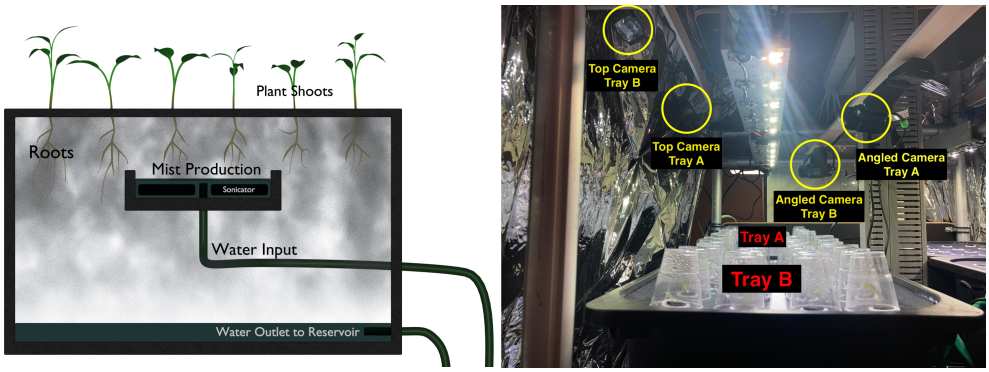

Figure 1. Growth bed schematic from the side (**left**). Camera rig for image capture (**right**).

3.2. Plant Physiology

The research conducted at the Swedish University of Agricultural Sciences (SLU) during a long period from the 1970s to the millennium shift has created a broad and solid scientific framework for controlled plant growth and development and, thereby, highly efficient and productive plant cultivation. This research is well-documented in scientific papers, dissertations and plant development databases [8,17–21] and provides a solid ground for our hypotheses and assumptions to be tested within the scope of this work.

One of the main assumptions in this work is that there exists a strong correlation and coherence between the relative rates of growth for biomass and leaf area. This has been verified in different experiments and is documented in [8,20], where biomass growth rates have been determined by weighing plants, roots and leaves, and leaf area growth rates have been determined by measuring the projected area via a scanner/copier. This correlation provides a strong indication that image analysis should be able to estimate biomass well.

Another assumption is that the RGR is constant for the plants grown for the data collection. A paper by O. Hellgren and T. Ingestad [17] demonstrated that, in controlled cultivation experiments with a constant relative addition rate of nutrients, plants have a constant relative growth rate. To achieve this, nutrients and water are supplied with free access, which implies non-limiting addition rates. This should ensure that the plant's RGR is close to constant, meaning that each gram of plant increases its mass by a constant amount per day.

3.3. Target Data

The ground target data is the biomass for a given plant at a given time point. This is gathered by weighing the plants at two time points. At each measurement, every plant is

weighed in its plant holder. The measured weight of the plant holder is then subtracted. These measurements then need to be extrapolated to every time point in the dataset. This is conducted through the assumption of a constant RGR, which is defined as

$$RGR = \dot{w}/w = a = const \Rightarrow \dot{w} = a \times w,$$

where $w(t)$ is the biomass. Since the growth rate is assumed to be linearly dependent on the biomass, then the biomass follows an exponential function. The log biomass is, therefore, in theory at least, a first-order polynomial with respect to time. Fitting this line to our measurements gives us an approximation of the log-biomass curve, which can be extrapolated to every time point. Note that the assumption of a constant RGR in the training set does not necessitate that future estimates on other plants require a constant RGR. This assumption only facilitates more efficient data gathering.

In reality, however, the plants exhibit an environmental shift when they are introduced to the cultivation platform. This means that the RGR could be lower early in the growth period before stabilizing later on to a constant value. The target data might, therefore initially differ slightly from the true biomass; however, they should eventually produce a reasonable approximation.

3.4. Input Images

The input data to the models consists of images of a specific plant at a given time during growth. These images were taken at 1 h intervals. At each such time point, a total of eight images are captured based on our two camera angles and the four growth beds. Each image covers an entire growth bed. These images are then transformed and grid-aligned through a projective transformation, shown in Figure 2, and then divided into segments, capturing a square around each plant. These images were resized to 64 × 64 pixels.

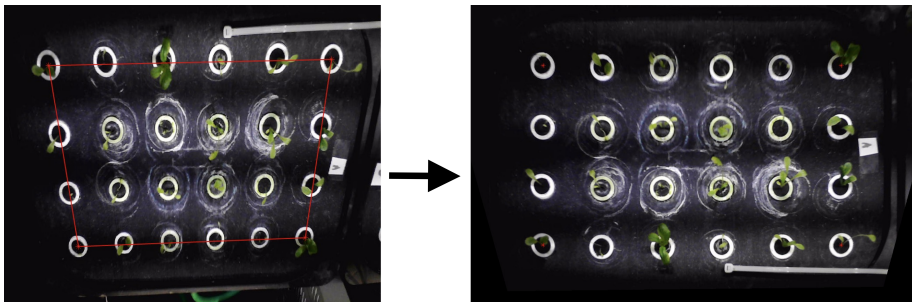

Figure 2. Full-tray image transformation and alignment.

In total , the dataset contains images of 57 plant individuals taken over a 5 day period, captured at 1 h intervals. This resulted in a total of 10,197 individual plant images. This dataset was split into training, validation and test sets. This split was performed based on individual plants, such that six randomly chosen individuals were placed in the validation and test sets, while the remaining 45 individuals were placed in the training set. This resulted in a [77.4%, 11.3% and 11.3%] split of the dataset. The dataset was further split into images from the top and angled camera respectively, resulting in the two datasets needed to train the two models for each ML method.

The dataset was pre-processed in a number of ways. First, the large degree of redundancy in the images increases the risk of over-training. This was combated through image augmentation by rotating each image by a random amount. A number of color spaces [22,23] were also however, the normal RGB color space was found to be optimal likely due to the pre-trained network having been trained on RGB images. The target biomass was also pre-processed through log transformation as this transformed the exponential biomass growth to a linear correlation. This led to the biomass being more evenly distributed. In addition, the log-biomass targets were normalized to the interval [0, 1]

based on the training set. The effect of these actions on the biomass distribution can be seen in Figure 3.

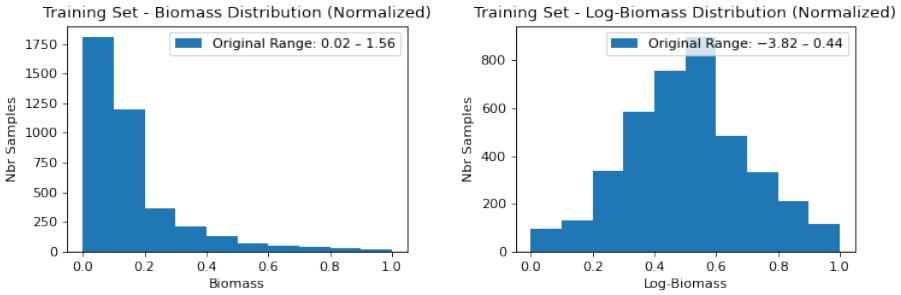

Figure 3. Effect of log-transform and normalization on biomass distribution for target data.

3.5. Method 1: Multi-Variate Regression

The first method used linear multi-variate regression on a number of features. The features used were all pixel-wise features summed over the image, meaning they all have the form

$$F_i(I) = \sum_{x=1}^{64} \sum_{y=1}^{64} f_i(I_{x,y}), \qquad f_i(p) : \mathbb{R}^3 \to \mathbb{R},$$

where I is an image and $I_{x,y}$ is a pixel in that image. The features used were inspired in part by features commonly used in plant segmentation [24,25]. The investigated pixel-wise features for the function f_i are shown in Table 1.

Table 1. Description of the 13 investigated features for the MVR model.

Set Name	Features	Description
Color	3	Amount of red, green or blue in pixel
Ratio	3	Ratio between R/G, G/B, B/R in pixel
Hue	3	Number of pixels with hue between R-G, G-B, B-R
Highlight	3	Same as 'Hue', but weighed by saturation
Vegetative Index [26]	1	$\text{Green} \times \text{Red}^{-2/3} \times \text{Blue}^{-1/3}$

To identify useful features, an iterative step-wise construction was employed. This was conducted by starting with an empty model containing only the intercept. All neighboring models, reached by adding or removing a feature, were evaluated, and the best model was chosen. This continued until the current model outperformed all of its neighbors. The models were compared using either the AIC (Akaike Information Criterion) or the Bayesian Information Criterion (BIC). Models were also created with respect to either targeting the biomass or the log biomass. This resulted in four models per camera view with respect to evaluation metric and a target unit. The models with the lowest RMSE on the validation set were chosen as the final model for each camera.

3.6. Method 2: R-50-Based Neural Network

The second method used is a convolutional neural network (CNN), inspired by previous research in biomass prediction [11]. The network utilizes a pre-trained image recognition network called ResNet-50 [27], trained on data from ImageNet [28], as a base followed by a custom-made regression head.

The R-50-based network relies on a residual block, which, in contrast to regular NN layers, in the left side of Figure 4, aims to predict the residual between the input to an imagined desired output. This is performed by adding the input to the residual block to the output as can be seen in the right side of Figure 4. This architecture allows for a very

deep NN without exhibiting the accuracy degradation that has been previously observed in such networks [27].

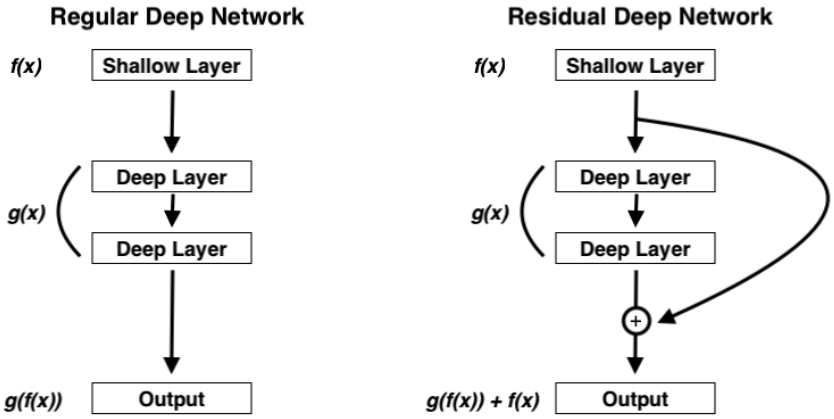

Figure 4. Architecture of a regular deep network (**left**) versus a residual deep network (**right**).

More specifically the R-50-based network includes a ResNet-50 network, which returns a feature vector of size 1000. This is then fed into a regression head, consisting of three densely connected layers of sizes 512, 128, and 1 respectively. The first two use ReLU activation, while the final output layer uses linear activation. Each network was trained for 20 epochs with a learning rate of 0.0005 and L2 regularization with a factor of 0.1. The pre-trained weights of the R-50-based network were also set to be trainable, so as to be fine-tuned for our problem. The architecture of the full network can be seen in Figure 5. Two models were created using images from the top and angled cameras respectively.

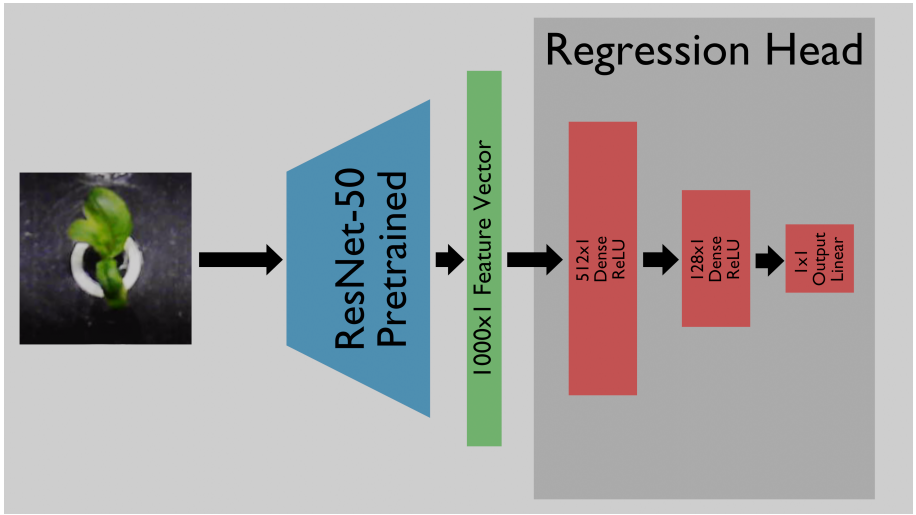

Figure 5. Graph of the architecture used for the R-50 network, including layer sizes, connection types, and activation function.

4. Results

In this section, we provide a summary of our main findings, which include training and testing results from the MVR and R-50-based network models as well as error comparisons of the different tasks depending on the number and type of image views undertaken.

4.1. Model Creation

It was found that AIC outperformed BIC on the validation set when constructing the MVR models. In addition, targeting the log biomass resulted in a higher accuracy on the validation set when using images from the top camera. Additionally, targeting

untransformed biomass was found to be superior for the angled camera. The parameter weights used in the final models for the top and angled camera are shown in Table 2.

Table 2. Features included in the two final MVR models.

Feature	Top View	Angled View
Intercept	0.1427	0.4567
R	−30.8498	−4.0938
G	29.9778	3.9349
B	-	-
R/G	−4.4577	−0.8282
G/B	−0.4435	1.0791
B/R	−4.3437	−0.7736
Hue_{RG}	0.8625	0.6988
Hue_{GB}	1.1628	0.4656
Hue_{BR}	-	-
HL_{RG}	−1.3134	-
HL_{GB}	−2.0212	-
HL_{BR}	6.2662	1.0448
VEG	-	−0.6336

Figure 6 shows the training process of the R-50-based network in terms of loss (MSE) for the top and angled cameras respectively.

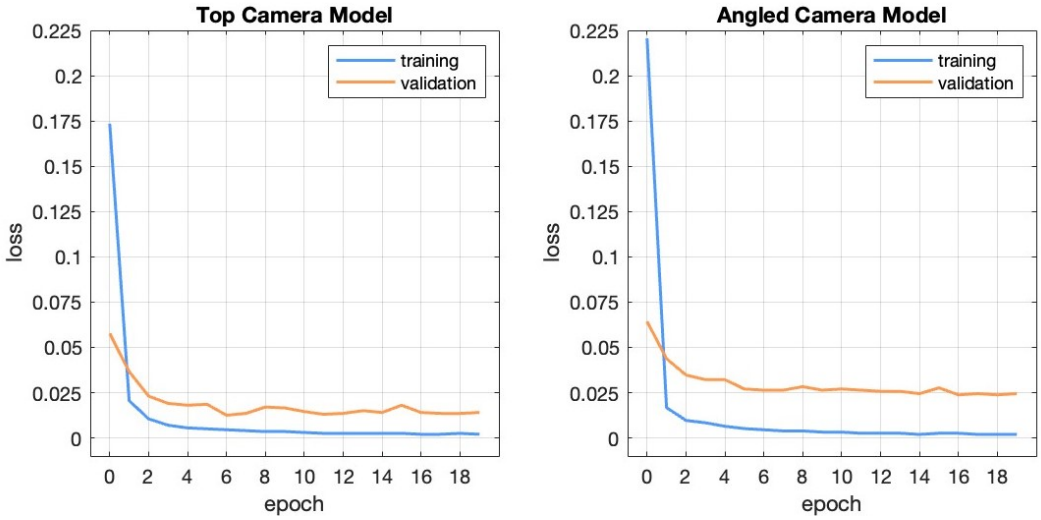

Figure 6. Training Loss (MSE) for the R-50 based network trained on images from the top (**left**) and angled (**right**) cameras respectively.

4.2. Method Comparison

Table 3 shows the RMSE on the test set for both the MVR network and R-50-based network. This table shows the quality of the method on a single image (SI) biomass prediction, the average (MA) of three chronologically consecutive biomass predictions, and RGR prediction using three random data points. Each of these tasks was performed using the model for the top view, angled view, and the average between their outputs (dual view). In addition, the RMSE is also presented in that table for the full test set as well as for some individual plants randomly chosen from the test set. These values are in the unit of g for biomass estimates and g/(g· day) for RGR estimates. For comparison, the true biomasses in the test set range up to 0.35 g and the true RGR range up to 0.17 g/(g· day).

Similarly, confidence intervals of the MSE on the test set with respect to camera view (Top, Angled, and Dual) and method (MVR and R-50 based network) are presented in Figure 7.

Table 3. RMSE on test set as a whole and a few, randomly picked, individual plants from the MVR network (above) and R-50 based network (below). Results are presented for all tasks: SI = Single Image, MA = Moving Average, RGR = Relative Growth Rate. The best method for a given task for each individual is highlighted in bold.

Multi-Variate Regression Method							
Task—View	All	#24	#33	#42	#64	#65	#78
SI—Top	**0.0466**	**0.0316**	**0.0260**	**0.0448**	**0.0573**	**0.0643**	0.0447
SI—Angled	0.1227	0.0924	0.0640	0.1347	0.2173	0.0960	0.0585
SI—Dual	0.0734	0.0521	0.0406	0.0882	0.1151	0.0751	**0.0363**
MA—Top	**0.0391**	**0.0253**	**0.0200**	**0.0426**	**0.0382**	**0.0614**	0.0332
MA—Angled	0.1178	0.0875	0.0598	0.1336	0.2089	0.0922	0.0479
MA—Dual	0.0703	0.0491	0.0380	0.0875	0.1096	0.0729	**0.0287**
RGR—Top	0.2268	0.2088	0.2038	0.1911	0.2413	0.2573	0.2500
RGR—Angled	0.3669	0.2816	0.2684	0.2261	0.2889	0.5410	0.4788
RGR—Dual	**0.1984**	**0.2049**	**0.1736**	**0.1658**	**0.2019**	**0.2383**	**0.1974**
ResNet-50 Method							
Task—View	All	#24	#33	#42	#64	#65	#78
SI—Top	0.0862	0.0339	0.0301	0.0261	0.1833	**0.0553**	**0.0744**
SI—Angled	0.0653	0.0265	0.0332	0.0284	0.0823	0.0809	0.0992
SI—Dual	**0.0550**	**0.0260**	**0.0228**	**0.0233**	**0.0735**	0.0639	0.0840
MA—Top	0.0812	0.0291	0.0278	**0.0166**	0.1752	**0.0502**	**0.0689**
MA—Angled	0.0637	**0.0214**	0.0304	0.0271	0.0804	0.0792	0.0985
MA—Dual	**0.0523**	0.0230	**0.0204**	0.0205	**0.0669**	0.0620	0.0829
RGR—Top	0.2647	0.2195	0.1982	0.3260	0.2444	0.2483	0.3243
RGR—Angled	0.1837	0.1910	0.1490	**0.1314**	**0.1818**	0.2294	0.2022
RGR—Dual	**0.1767**	**0.1434**	**0.1427**	0.1687	0.1867	**0.2078**	**0.2000**

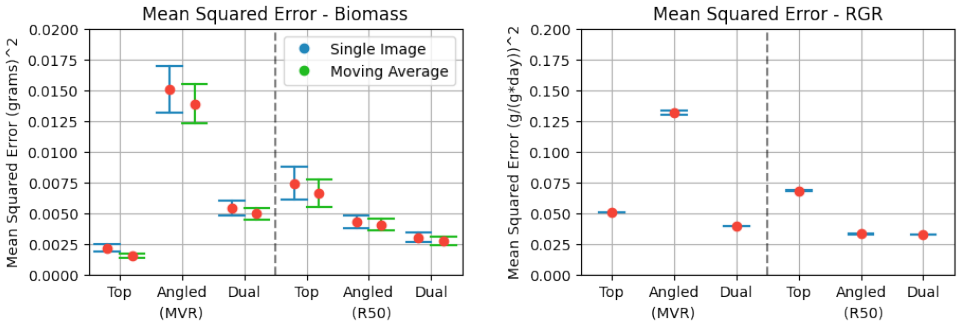

Figure 7. Confidence intervals (95%) for mean square errors in prediction on the test set using MVR or R-50 based network, with respect to camera view (top, angled, or dual). (**Left**): Biomass predictions in tasks Single Image and Moving Average. (**Right**): RGR predictions in task RGR.

5. Discussion

We trained a multi-variate regression and a R-50-based convolutional neural network, on a combination of images from different viewpoints and observed their performance towards estimating plant biomass and relative growth rate.

5.1. Multi-Variate Regression

It was found that the Akaike Information Criterion (AIC) metric outperformed the Bayesian Information Criterion (BIC) metric when generating MVR models. In [29], a comparison was made between AIC and BIC and they concluded that BIC had an advantage over AIC if the model used to generate the dataset was included in the set of possible

models. They also noteed, however, that, since data from the real world are too complicated, "the primary foundations of the BIC criteria do not apply in the biological sciences and medicine and the other 'noisy' sciences" [29]. This also holds true for our case, since it is reasonable to believe that the biomass of the plants was not generated using the pixel values of the captured images. This could be the reason that the AIC performs much better.

Overall, MVR performed substantially better on data from the top camera compared to the angled camera data as can be seen in Figure 7. This holds true on an individual level as can be seen in rows 1–2 in Table 3. This is reasonable since the top camera is a better representation of the leaf area and is thus a better representation of the biomass. In addition, the nature of the angled camera causes parts of other plants to be included in the images. Since MVR has no way of distinguishing between these pixels, there is an inherent flaw with applying this non-structural method to the images from the angled camera.

Similar to the biomass estimate, the RGR estimate was significantly better for the top camera compared to the angled camera. However, for the RGR estimates, the dual view was found to be superior. This shows that a poor estimate of biomass did not necessarily lead to a poor RGR estimate, even on an individual level. The selection of the three random samples could, on the other hand, be affected by outliers. This might cause the variance-reducing effect of the dual view to be more prominent in the RGR estimate than in the biomass estimate, thus leading to an improvement in the former, but not the latter.

5.2. The ResNet-50-Based Convolutional Neural Network

For the R-50-based network, the top camera underperformed than the angled camera as seen in Figure 7. However, this only holds true when looking at the overall RMSE. When investigating the RMSE on the individual plants in the lower section of Table 3, the top camera performed better on four out of six plants. Plant #64 had a surprisingly large RMSE in the top view, which inflates the overall RMSE.

As opposed to the results from the MVR, the dual view performed better than both the top and angled views. This indicates that having two perspectives is beneficial for the R-50-based network. The reason for this could be that the difference in MAE between the top and angled view is much smaller, making the variance-reducing effect of their average more prominent.

Another pattern that was found in the R-50-based network results is that the RGR estimate was better when using the dual view compared to the top or angled view.

5.3. Method Comparison

Comparing the RMSE between the two methods shows that the MVR performed best on the top camera, while the R-50-based network performed the best on the angled camera as can be seen in Figure 7. As mentioned, the reason that the MVR has a poor result on the angled view could be that a smaller fraction of the image consists of plant pixels, and that other plants can show up and distort the estimate. However, the additional structural information in the angled view, such as plant height, could be used by the R-50-based model which might be why it performs better.

The dual view inhibits some interesting behaviors. Since the dual view is created from the average of the top and angled camera, the performance is somewhat close to their average. However, it is always slightly lower. The reason for this is likely that the averaging also has a variance-reducing effect, leading to better estimates. We can see these effects in the left graph of Figure 7. This means that the dual view is superior when the difference between the top and angled cameras is small, such as in the R-50-based method. But if the difference is large, such as in the MVR method, the variance-reducing effect is not large enough to overcome the benefits of using only the superior camera.

In every case, the moving average estimate had a lower RMSE than the corresponding RMSE for the single-image task. This shows that there is a benefit to capturing multiple images at a high frequency, even though the biomass increase is minimal. This benefit likely comes from the fact that the flickering of the LED lights was noticeable in the images.

The noise from this flickering likely leads to some variation in the estimates, which is reduced by the moving average filter.

In the case of estimating the RGR, we found that, in general, the dual view was found to be superior for both methods.

5.4. Prediction Quality

The best result for the MVR model produced an RMSE of around 0.0391 g using the moving average filter and top camera images. With biomasses up to 0.35 g in our test set, this represents a relative RMSE of 11.2%, which is comparable to previous studies. In the paper by Wenjian Liu et al. [24], for instance, they achieved an RMSE of 0.32 g on fresh biomass samples up to 3 g, which corresponds to a relative RMSE of 10.7%. Other papers have achieved even better accuracy for datasets with larger plant biomass weights. The paper by N. Buxbaum et al. [11], for instance, used images of 3888 individuals and obtained an RMSE of between 1 and 2 g for plant biomass up to 40 g, corresponding to a relative RMSE of 2.5–5%.

We believe that a reason for this value of relative RMSE could be attributed to the large variance in the estimates of time-adjacent images due to the flicker of the LED lights creating visible variations between the images. Having a setup where the collected images are not affected by such external factors should therefore improve both the biomass as well as the RGR estimates.

There do not exist any aeroponic studies to compare these results against, but in general, the RGR estimates were not sufficiently accurate. Detecting relative changes in biomass could be easier for larger plants, as the visual difference is larger the further into the growth period the plants are. A dataset containing more individuals would likely also increase the accuracy of the predictions. It should be noted that the RGR estimates were made using three random points from the entire growth period. The resulting estimate could be considered to measure the 'average' RGR. Since our data were assumed to have a constant RGR this does not matter. However, for conditions where the RGR varies over time (for example, if the conditions change during growth), the RGR should instead be constructed from a time series with a shorter time span depending on the time resolution desired.

6. Conclusions

In this work, we trained two different machine-learning models, a multi-variate regression model (MVR) and a ResNet-50 (R-50) neural network, to discover growth patterns in plants based solely on camera images. We then compared the abilitys of those models to forecast plant biomass and plant relative growth rate (RGR). Our proposed approach to estimate plant development, therefore, relies on non-destructive methods. As a result, it can be possible for farmers to intervene at an early stage if needed, in order to influence and improve growth even at the individual plant level.

We note also that, in general for any type of plant growing environment, as long as data is recorded consistently the resulting model will be able to learn. In our study, we only required a short series of images in order to estimate biomass. In that respect, changes in atmospheric or soil moisture did not influence the imaging of the plants which in turn are responsible for the accuracy of our estimates.

Based on the results in Section 5 above, we see that the biomass estimate can be improved greatly for both models when applying the moving average filter over the neighboring time points. Having multiple cameras does not seem to improve the estimates from the multi-variate regression model, but can improve the estimates based on the neural network model.

We also see that the best RGR estimates can be produced when images from both cameras are used to train the neural network model. More generally however, using images from both cameras improves the RGR estimate for both the multi-variate regression as well as the R-50 based neural network.

Although the quality of the resulting biomass predictions is comparable to other studies, the way these predictions are produced is significantly different. They are based on non-destructive data collection, and as a result, they can only be improved as more data become available for the study. There currently do not exist any other similar results investigating the ability of machine learning to predict RGR or biomass from images in aeroponics.

The findings in this study highlight promising patterns in camera and model behavior, such as the effects of moving average filters and multiple camera angles. Such approaches also outline the potential for future research into virtual sensors of RGR. Such sensors would contain the cameras as well as a computer analyzing the images using the models and transmitting the estimated biomass and RGR. Further research in this field, using datasets of more individuals over long periods of time and using more real measurements, is vital in order to verify these results.

Author Contributions: Conceptualization, O.Å. and H.H.; Formal analysis, O.Å., H.H. and A.S.; Funding acquisition, H.H. and A.S.; Investigation, O.Å., H.H. and A.S.; Methodology, O.Å., H.H. and A.S.; Project administration, O.Å., H.H. and A.S.; Resources, O.Å., H.H. and A.S.; Software, O.Å. and A.S.; Supervision, H.H. and A.S.; Validation, O.Å., H.H. and A.S.; Visualization, O.Å.; Writing—original draft, O.Å., H.H. and A.S.; Writing—review and editing, O.Å. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research is partially supported by grants from eSSENCE (number 138227), Vinnova (number 2020-03375), Formas 2022-00757, and the Swedish National Space Board.

Data Availability Statement: Data available on request due to restrictions eg privacy or ethical. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to persons in the background of the side images.

Acknowledgments: The training and data handling was enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC), partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

Conflicts of Interest: The authors declare no conflict of interest. Some of the data in this study were provided by Alovivum AB. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ML	Machine learning
RGR	Relative growth rate
MVR	Multi-variate regression
R-50	Neural network with ResNet-50 base
NN	Neural network
CNN	Convolutional neural network
MSE	Mean square error
RMSE	Residual mean square error

References

1. Olympios, C.M. Overview of soilless culture: Advantages, constraints and perspectives for its use in Mediterranean countries. *Cah. Options Méditerranéennes* **1999**, *1*, 307–324.
2. Ghorbel, R.; Chakchak, J.; Malayoğlu, H.; Çetin, N. Hydroponics "Soilless Farming": The Future of Food and Agriculture—A Review. In Proceedings of the 5th International Students Science CongressProceedings, Rome, Italy, 20–22 October 2021.
3. Sheikh, B.A. Hydroponics: Key to sustain agriculture in water stressed and urban environment. *Pak. J. Agric. Agric. Eng. Vet. Sci.* **2006**, *22*, 53–57.
4. Tunio, M.H.; Gao, J.; Shaikh, S.A.; Lakhiar, I.A.; Qureshi, W.A.; Solangi1, K.A.; Chandio, F.A. Potato production in aeroponics: An emerging food growing system in sustainable agriculture forfood security. *Chil. J. Agric. Res.* **2020**, *80*, 118–132. [CrossRef]
5. Ziegler, R. *The Vertical Aeroponic Growing System*; Synergy International Inc.: Sausalito, CA, USA, 2015.

6. Mokhtar, A.; El-Ssawy, W.; He, H.; Al-Anasari, N.; Sammen, S.S.; Gyasi-Agyei, Y.; Abuarab, M. Using Machine Learning Models to Predict Hydroponically Grown Lettuce Yield. *Front. Plant Sci.* **2022**, *13*, 706042. [CrossRef]
7. Gao, D.; Qiao, L.; An, L.; Zhao, R.; Sun, H.; Li, M.; Tang, W.; Wang, N. Estimation of spectral responses and chlorophyll based on growth stage effects explored by machine learning methods. *Crop J.* **2022**, *10*, 1292–1302.
8. Hedlund, H. *Temperature Distribution and Plant Responses of Birch (Betula Pendula Roth.) at Constant Growth*; Acta Universitatis Agriculturae Sueciae Agraria, Swedish University of Agricultural Sciences: Uppsala, Sweden, 1999.
9. Carter, W.A. A method of growing plants in water vapor to facilitate examination of roots. *Phytopathology* **1942**, *732*, 623–625.
10. Ojo, M.O.; Zahid, A. Deep Learning in Controlled Environment Agriculture: A Review of Recent Advancements, Challenges and Prospects. *Sensors* **2022**, *22*, 7965. [CrossRef]
11. Buxbaum, N.; Lieth, J.H.; Earles, M. Non-destructive Plant Biomass Monitoring With High Spatio-Temporal Resolution via Proximal RGB-D Imagery and End-to-End Deep Learning. *Front. Plant Sci.* **2022**, *13*, 758818. [CrossRef]
12. Jung, D.H.E.A. Image Processing Methods for Measurement of Lettuce Fresh Weight. *J. Biosyst. Eng.* **2015**, *40*, 89–93. [CrossRef]
13. Beck, M.A.; Liu, C.; Bidinosti, C.P.; Henry, C.J.; Godee, C.M.; Ajmani, M. Presenting an extensive lab- and field-image dataset of crops and weeds for computer vision tasks in agriculture. *arXiv* **2021**, arXiv:2108.05789.
14. Mehra, M.; Saxena, S.; Sankaranarayanan, S.; Tom, R.J.; Veeramanikandan, M. IoT based hydroponics system using Deep Neural Networks. *Comput. Electron. Agric.* **2018**, *155*, 473–486. [CrossRef]
15. Broms, C.; Nilsson, M.; Oxenstierna, A.; Sopasakis, A.; Åström, K. Combined analysis of satellite and ground data for winter wheat yield forecasting. *Smart Agric. Technol.* **2023**, *3*, 100107. [CrossRef]
16. Kumar, P.; Prasad, R.; Gupta, D.K.; Mishra, V.N.; Vishwakarma, A.K.; Yadav, V.P.; Bala, R.; Choudhary, A.; Avtar, R. Estimation of winter wheat crop growth parameters using time series Sentinel-1A SAR data. *Geocarto Int.* **2018**, *33*, 942–956.
17. Hellgren, O.; Ingestad, T. A comparison between methods used to control nutrient supply. *J. Exp. Bot.* **1996**, *47*, 117–122. [CrossRef]
18. Ingestad, T.; Hellgren, O.; Lund Ingestad, A. *Data Base for Birch Plants at Steady State*; Technical Report 75; Sveriges Lantbruksuniversitet Rapporter: Uppsala, Sweden, 1994 .
19. Hellgren, O.; Ingestad, T.; Lund Ingestad, A. *Data Base for Tomato Plants at Steady-State—Methods and Performance of Tomato Plants (Lycopersicon esculentum Mill cv Solentos) under Non-Limiting Conditions and under Limitation of Nitrogen and Light*; Technical Report 74; Institutionen foer Ekologi och Miljoevaard (Sweden): Uppsala, Sweden, 1994 .
20. Hellgren, O.; Ingestad, T. *Responses of Birch (Betula Pendula Roth) and Tomato Plants (Lycopersicon Esculentum Mill cv Solentos) to CO_2 Concentration and to Limiting and Non-Limiting Supply of CO_2*; Technical Report 3; Biotron, Swedish University of Agricultural Sciences: Uppsala, Sweden, 1996.
21. McDonald, A.J.S.; Lohammar, T.; Ingestad, T. Net assimilation rate and shoot area development in birch (Betula pendula Roth.) at different steady-state values of nutrition and photon flux density. *Trees* **1992**, *6*, 1–6. [CrossRef]
22. Praveen Kumar, J.; Domnic, S. Image based leaf segmentation and counting in rosette plants. *Inf. Process. Agric.* **2019**, *6*, 233–246. [CrossRef]
23. Yang, W.; Wang, S.; Zhao, X.; Zhang, J.; Feng, J. Greenness identification based on HSV decision tree. *Inf. Process. Agric.* **2015**, *2*, 149–160. [CrossRef]
24. Liu, W.; Li, Y.; Liu, J.; Jiang, J. Estimation of Plant Height and Aboveground Biomass of Toona sinensis under Drought Stress Using RGB-D Imaging. *Forests* **2021**, *12*, 1747. [CrossRef]
25. Lati, R.N.; Filin, S.; Eizenberg, H. Robust Methods for Measurement of Leaf-Cover Area and Biomass from Image Data. *Weed Sci.* **2011**, *59*, 276–284. [CrossRef]
26. Hague, T.; Tillett, N.D.; Wheeler, H.C. Automated Crop and Weed Monitoring in Widely Spaced Cereals. *Precis. Agric.* **2006**, *7*, 21–32. [CrossRef]
27. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)* **2016**, *322*, 770–778. [CrossRef]
28. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
29. Burnham, K.; Anderson, D. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*; Springer: Berlin/Heidelberg, Germany, 2002.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.