HIGH SCHOOL EDITION
Journal of Student Research

Volume 12 Issue 1 (2023)

Fully Automatic Controlled Environment Agriculture Using Machine Learning Based Plant Size Estimator

Seung Wook Yoon[1] and Sukyeong Nam[#]

[1]Hankuk Academy of Foreign Studies, Republic of Korea
[#]Advisor

ABSTRACT

Due to climate change, the global food supply crisis has become an urgent international problem. Given the circumstances, a controlled environment that artificially adjusts the climate for agriculture has attracted considerable attention as a solution to the problem. Various kinds of research have been proposed to develop the technology of controlled environments. However, the accuracy and scalability problem of these methods is a burden for the expansion to real-world scenarios. In particular, it is necessary to research and implement the computer vision-based algorithm, which is the key technique that enables the controlled environment system to be fully automatic. To solve the aforementioned problem, I propose a novel controlled environment agriculture system. The proposed system is composed of a plant life cycle regression module and a device control module. The system predicts the actual size of the plants and outputs the life cycle indicator of the plant, which is the growth rate of the plants. Based on the life cycle indicator, the device control module adjusts the essential factors, such as the amount of water and the strength of UV (Ultra Violet) light, for photosynthesis. As the proposed system is aware of the life cycle of plants, it can provide fully automatic controlled environments. I also propose and demonstrate the application machine to show how the proposed method can be applied to the real world. Through the experiments, it is shown that the proposed PLCR outperforms the existing state-of-the-art methods on the COCO dataset.

Introduction

As the environment on Earth becomes less suitable for agriculture due to pollution or climate change, the food supply crisis is a critical issue that humans should deal with (Robinson et al. 2001). Unpredictable global problems such as the Ukraine war complicates the problem of the whole food supply chain, exacerbating the food supply crisis as well. To find the answer to this problem, researchers have been searching for technology that can efficiently increase the food supply without the environmental constraints of farming. For example, certain crops are only produced in precise humidity, soil pH, and precipitation. This is a substantial constraint for some of the countries in crisis.

Agricultural technologies have been widely explored to solve such problems. Traditionally, researchers have been developing a controlled environment, which is the system that controls the factors that engage in the growth of a plant, such as water, temperature, or humidity (Castelló et al. 2018). The researchers have focused on implementing the system interface to easily control the aforementioned factors as intended. Their method reduced the demanded labor power to maintain the quality of products, but the decision or the action of humans is still required.

Recently, to overcome this limitation, a few studies have developed automatic controlled-environment systems that do not need human interference (Menon et al. 2021).

Their methods are based on deep learning technologies that have shown remarkable performance in many vision domains. However, they have scalability issues in that it is practically impossible to expand the system to real-world scenarios. To solve the scalability issue, it is necessary to develop the system with human-level precision to make the system reliable.

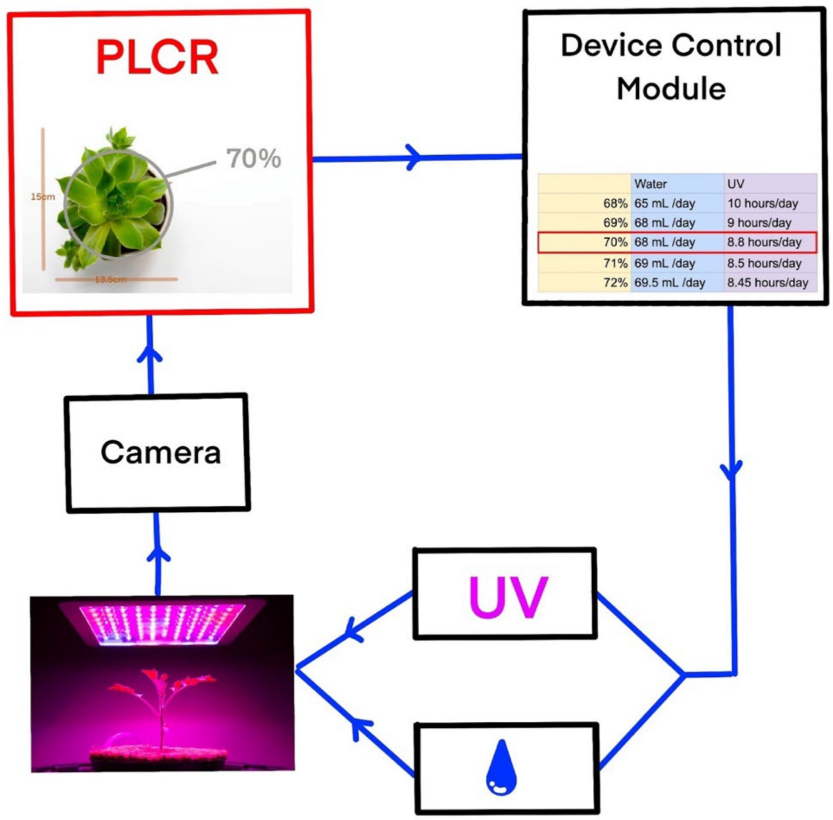


Figure 1. Flow chart of the proposed controlled environment agriculture system. (The proposed PLCR takes the plant image obtained via the top-view installed camera and outputs the life cycle indicator of the plant. Based on the predicted life cycle indicator, the embedded water pump and light source are operated by the device control module.)

In this paper, I propose a novel controlled environment agriculture system. The proposed system is composed of a device control module, and a plant life cycle regression module, as shown in Figure 1. The device control module contains the water circulation pump, which keeps the consistent water level in the water tank based on the climate recipe. Furthermore, the model controls a UV light source, providing artificial plant sunlight. In conclusion, the device control module controls and adjusts the factors that engage in the growth of a plant. The proposed system outputs the life cycle indicator of the plant by comparing the measurement with the size of full-grown plants. To accomplish this, I design the regression model that predicts the bounding box of the plants using convolutional neural networks. I also propose a size conversion method to measure the actual size of plants and their growth rate. Overall, the proposed method enables the full automation of controlled environment agriculture systems.

Related Work

Controlled Environment

A controlled environment engages with precisely regulated environmental factors for plants to grow, such as humidity, sunlight, or water. A controlled Environment can be used to harvest foods that are conventionally not growable in a particular region. For example, tropical fruits usually grown in tropical regions can be grown in any area as long as a controlled environment is set up. As global warming and climate change caused harvest crises, leading to danger to the food supply chain, the controlled environment emerged as the key to this problem. MIT (Massachusetts Institute of Technology) media laboratory proposed a PFC (Personal Food Computer) (Castelló et al. 2018), which is a device for the management of a controlled environment system. The PFC can control the essential factors for the plants' photosynthesis, such as temperature or water level. This method has shown that it can grow foods in a controlled environment adjusted by a computer. However, these controls have the disadvantage that human intervention is necessary.

Inspired by previous research, I started researching the controlled environment of real life. I propose a novel controlled environment agriculture system in this paper. I aimed to propose a system that is aware of the life cycle of plants based on machine learning technology so that it can provide fully automatic controlled environments.

Object Detection

Object detection predicts the category and location of the objects in images or videos. The categorization process is conducted the same as image classification. Object localization is often performed via predicting the bounding box that surrounds the object.

Object detection can be categorized into one-stage and two-stage methods. One-stage methods usually operate faster than the two-stage methods as they combine the region proposal and region rejection steps. SSD (Liu et al. 2016) or YOLO (Redmon, et al. 2016) are representative one-stage methods. For two-stage methods, Faster R-CNN (Ren, et al. 2015) is often used since it tends to produce more accurate results than the one-stage methods. In this paper, I consider the plant size estimator as an object localization process. The proposed system predicts the bounding box of the plant in the inputted image without a categorization score. The detailed process is further explained in chapter 3.

Methods

Figure 1 represents the overall architecture of the proposed system. The proposed system is composed of a PLCR (Plant Life Cycle Regression) module and a device control module. The PLCR module inputs a top-view image of the plant and outputs the life cycle indicator of the plant, which is represented as a growth rate. As shown in Figure 2, the PLCR module measures the actual size of the plants and outputs the life cycle indicator of the plant by comparing the measurement with the size of full-grown plants.

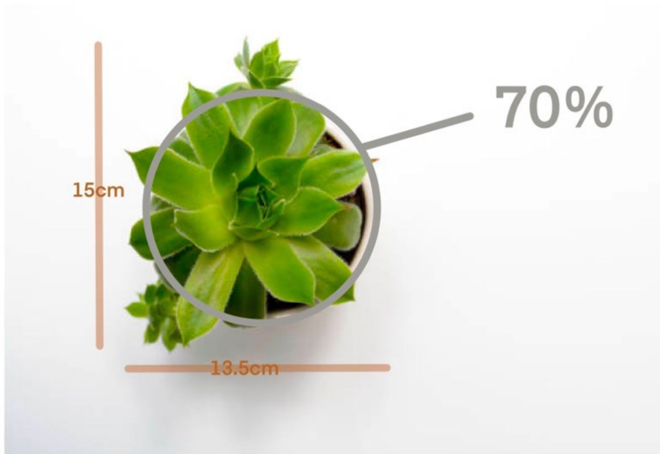


Figure 2. Example of the output of the proposed PLCR (Plant Life Cycle Regression) module.

Based on the output of the PLCR, the device control module adjusts the pre-installed devices to control the essential factors, which are the amount of water and the strength of UV light for the plants' photosynthesis. Plants require adaptive essential factors at a particular growth rate to make production more effective. For example, some plants require more water and sunlight in the early growth stage; others need more water and sunlight when they mature. The device control module can automatically adjust the essential factors for the plants with the estimated life cycle indicator from the PLCR.

PLCR

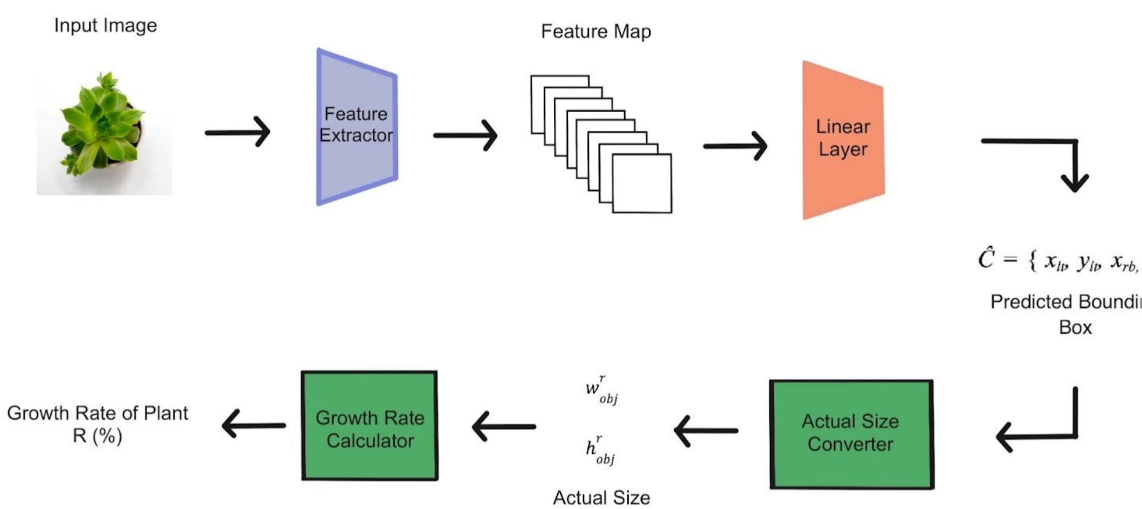


Figure 3. The architecture of the proposed PLCR (Plant Life Cycle Regression) module

Figure 3 shows the architecture of the proposed PLCR module. It is composed of object localization and size convert modules. The object localization module is developed based on convolutional neural networks. The given top-view plant image is fed to the feature extractor and converted into a feature map. The feature map is then flattened and fed to the two linear layers to predict the coordinates of the bounding box of the plant in the

HIGH SCHOOL EDITION
Journal of Student Research

Volume 12 Issue 1 (2023)

image. I define the predicted bounding box as set $\hat{C} = \{ x_{lt}, y_{lt}, x_{rb}, y_{rb} \}$, containing each coordinate value of the bounding box. Here, x_{lt}, y_{lt}, x_{rb}, and y_{rb} denote the x coordinate of the left top point, the y coordinate of the left top point, the x coordinate of the right bottom point, and the y coordinate of the right bottom point, respectively.

To measure the actual size of the plant, the pixel size of the estimated bounding box needs to be converted into an actual size unit. To accomplish this, I proposed a size-converting process by applying proportional relationships between the bounding box of the reference object and the plant. The detailed calculation process is defined as Equation 1.

Equation 1: Actual size converter:

$$w^p_{obj} : w^p_{ref} = w^r_{obj} : w^r_{ref}$$
$$h^p_{obj} : h^p_{ref} = h^r_{obj} : h^r_{ref}$$

Where, w^p and h^p denote the width and height of the bounding box in pixels, while w^r and h^r denote the actual width and height of an object in millimeters units. For example, h^p_{obj} is the height of the target objects in pixels and is the actual width of the reference object in millimeter units.

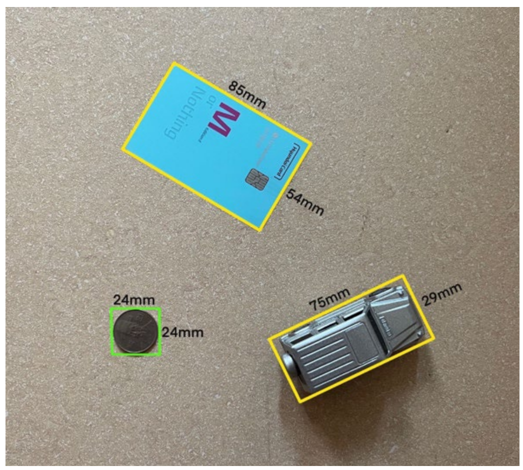


Figure 4. Example of size unit convert process using proportional relationships between the reference object(green) and the predicted bounding box(yellow)

I use a coin for a reference object as it has a rigid body and is geometrically symmetric, which makes the localization easier. As shown in figure 4, the pixel height and width of the predicted bounding box of a target object can be easily converted into actual size by using the proposed size conversion method.

Finally, PCLR outputs the growth rate of plant R based on the calculated actual size from the size conversion module. The growth rate R is calculated by dividing the area of the plant by the predefined area of the full-growth plant as Equation 2.

Equation 2: Growth rate equation

$$Plant_{area} = |x_{lt} - x_{rb}| \times |x_{lt} - x_{rb}|$$
$$R = (Plant_{area} / Target_{area}) \times 100$$

Where, $Plant_{area}$ and $Target_{area}$ denote the area of the quadrangle of the input plant and full-grown plant, respectively. After the calculation of the $Plant_{area}$, the growth rate of the plant R(%) is calculated. R represents how much the plant is grown. For example, if result R is 50%, the plant is only halfway grown compared to the well-grown plant. On the other hand, when the PLCR measures the R as 100%, we can harvest the plant.

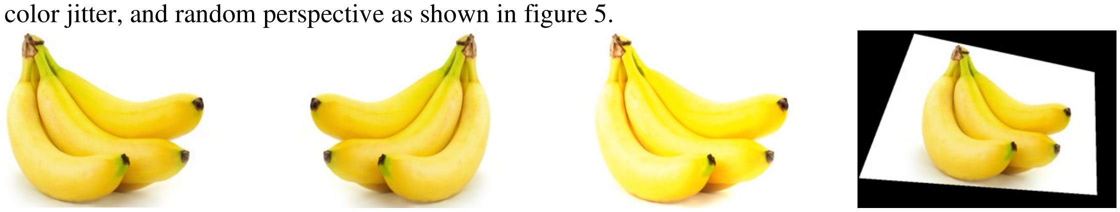
The loss function quantitatively measures how accurate the trained model's prediction is. To train the proposed object localization network, I use MSE (Mean Squared Error) loss function, which is often used to train localization networks (Huang et al. 2018), (Ren et al. 2015). The MSE loss function is calculated as Eq. (3).

Equation 3: MSE loss function:

$$L_{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{C}_i - C_i)^2$$

Here, $\hat{C}$ and C denote the predicted bounding box and its ground truth, and n is the number of the sample. The ideal model loss would be 0, meaning that the predicted bounding box is perfectly accurate.

For the training process of PLCR, we used the optimizer Adam. The initial learning rate was 0.0001, and I decayed the learning rate for every 80 epochs by 1/10. The total epoch was 200, so the decay was done at 80 and 160. Also, the batch size was 32. The data augmentation methods I used for the training were horizontal flip, color jitter, and random perspective as shown in figure 5.



(a) Original Image (b) Horizontal Flip (c) Color Jitter (d) Random Perspective

Figure 5. Example of data augmentation technique

Device Control Module

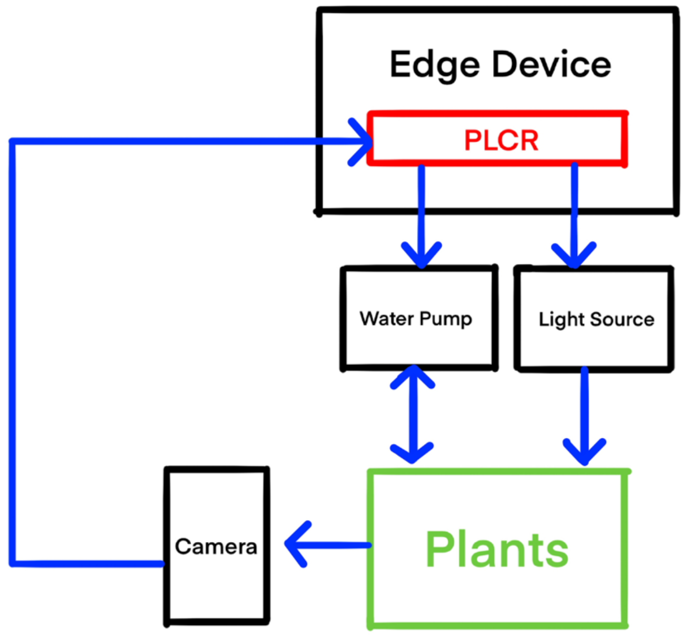


Figure 6. Diagram of the process of the device control module

Figure 6 is a diagram that describes the whole process of how the proposed PLCR adjusts the device control modules. The PLCR, which is implemented on the edge device, controls the overall mechanical process of the device control module. The proposed PLCR takes the image of plants as input and outputs the life cycle indicator of the plant. Based on the predicted life cycle indicator, the embedded water pump and light source are operated and scheduled for the photosynthesis of plants. For the computation process device, I choose Raspberry Pi, a single-board computer with low power consumption.

Experimental Results

Dataset

Figure 7. A snippet of COCO (Lin et al. 2014) dataset.

I used the dataset COCO (Lin et al. 2014) to train the proposed system. The dataset COCO, collected by Microsoft, is large-scale object detection, segmentation, and captioning dataset, and it contains more than 200,000 images with 80 categories that can be used for training. Figure 7 shows the example of datasets that are contained in the COCO. Specifically, I use the provided object detection annotation in this paper to train the object localization network.

Quantitative Evaluation

For the quantitative evaluation metric for the proposed object localization network, I used the IoU (Intersection over Union), which is often used for the object detection problem (Ren et al. 2015), (Liu et al. 2016).

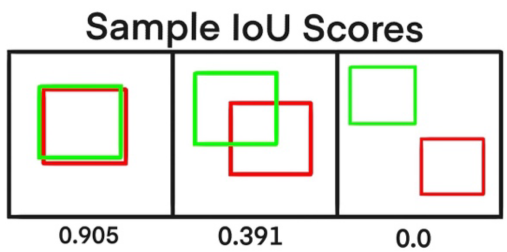


Figure 8. An example of IoU (Intersection over Union) evaluation metric. (Red box: ground truth, Green box: predicted bounding box)

IOU is the score value that is measured between 0 to 1. The overlap part of the answer and the predicted would be divided by the union area of both parts. Higher overlap with less union indicates that the predicted part is close to the answer. Therefore, 0 is the poorest score it can have, and 1 is the best score it can have. IoU value is calculated by Equation 4.

Equation 4: Example of an equation that can be cited later in the article text:

$$IoU = Plant_{area} \cap GroundTruth_{area} / Plant_{area} \cup GroundTruth_{area}$$

Where, $Plant_{area}$ is the prediction area of the PLCR, and $GroundTruth_{area}$ is the answer area in training. The area of the green box in Fig 8 is the $Plant_{area}$, and the red box is the $GroundTruth_{area}$.

Table 1. IoU evaluation comparison with state-of-the-art method

	IoU
SSD (Liu et al. 2016)	0.6972
Faster R-CNN (Ren et al. 2015)	0.7459
Ours	0.7871

Table 1 compares the IoU evaluation results with the state-of-the-art methods. For the comparison methods, I choose SSD (Liu et al. 2016) and Faster R-CNN (Ren et al. 2015), which generally show comparable performance in object detection problems.

The proposed method achieves an IoU of 0.7871, while the SSD and Faster R-CNN achieve an IoU of 0.6972 and 0.7459, respectively. Compared with SSD, the proposed method's score is 0.0899 higher, and compared with Faster R-CNN, the proposed method's score is 0.0412 higher. The proposed method extracts richer features compared to SSD and Faster R-CNN, leading to higher IoU results.

Table 2. Actual size error comparison

	Average error (width)	Average error (height)
SSD	16.7 cm	14.5 cm
Faster R-CNN	8.8 cm	10.1 cm
Ours	7.4 cm	8.2 cm

To examine the effectiveness of the proposed size estimation in a real-world scenario, I also conducted an additional experiment that can verify that the proposed method is more effective for measurement. I have divided the experiment into two categories: the error of width and length. I averaged the error of length and width of 30 photos of crops such as apples, cucumbers, bananas, and other plants.

The error of the proposed size estimation was 7.4 and 8.2 cm for the width and height. SSD has 16.7 cm in width and 14.5 cm in height. Faster R-CNN achieves 8.8 cm in width and 10.1 cm in height. The result shows that the proposed size estimation has a lower error compared to SSD and Faster R-CNN. Compared to SSD, the proposed method has a 9.3 cm lower width error and a 6.3 cm lower height error. Similarly, compared to Faster R-CNN, the proposed method achieves a 1.4 cm lower width error and 1.9 lower height error. This result shows that the proposed model is not only superior for the prediction of the bounding box but also for the real-life size estimation.

Ablation Study

In this chapter, I conduct an ablation study to investigate the effectiveness of each proposed idea. I measured the performance of the full model, which contained all the ideas, including the data augmentation process. As table.4 shows, the performance measured was 0.7871. Next, I measured the performance without the data augmentation process, and the result was 0.7608. As a result, performance without data augmentation was 3.34% lower than the full model, which means that the absence of data augmentation was more critical for the degeneration of the model performance.

Table 3. Ablation Study Result

Method	IoU
w/o data augmentation	0.7608
full model	0.7871

Conclusion

In this paper, to overcome the scalability issues of the current controlled environment agriculture system, I propose a novel controlled agriculture system, which is composed of a device control module and a PLCR(Plant Life Cycle Regression) module.

The device control module's role is to adjust the factors that engage in the growth of a plant, such as sunlights or water. PLCR's role is to measure the size of the plant and estimate the growth rate based on the measurement. I used dataset COCO for the training and testing of PLCR. For the evaluation, I used the IoU and a comparison between the existing models SSD and Faster R-CNN. I also progressed the ablation study to show how data augmentation contributes to the final performance of the trained model. The result showed that the proposed model is up to 12.8% superior to the existing state-of-the-art methods.

In this paper, to overcome the scalability issues of the current controlled environment agriculture system, I propose a novel controlled agriculture system, which is composed of a device control module and a PLCR(Plant Life Cycle Regression) module. The device control module's role is to adjust the factors that engage in the growth of a plant, such as sunlight or water. PLCR's role is to measure the size of the plant and estimate the growth rate based on the measurement. I used the COCO dataset for the training and testing of PLCR. For the evaluation, I used the IoU and a comparison between the existing models SSD and Faster R-CNN. I also progressed the ablation study to show how data augmentation contributes to the final performance of the trained model. The result showed that the proposed model is up to 12.8% superior to the existing state-of-the-art methods. In the future, based on the proposed method, I plan to develop an additional feature of the plant disease recognition system to expand the system to be more systematic in caring for the growth of plants by the system itself.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

Reference

Castelló Ferrer, Eduardo, et al. "Personal Food Computer: A new device for controlled-environment agriculture." *Proceedings of the Future Technologies Conference*. Springer, Cham, 2018.

He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

Huang, Rachel, Jonathan Pedoeem, and Cuixian Chen. "YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers." *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018.

Lin, Tsung-Yi, et al. "Microsoft coco: Common objects in context." *European conference on computer vision*. Springer, Cham, 2014.

Liu, Wei, et al. "Ssd: Single shot multibox detector." *European conference on computer vision*. Springer, Cham, 2016.

Menon, Vishal, V. Ashwin, and Raj K. Deepa. "Plant Disease Detection using CNN and Transfer Learning." *2021 International Conference on Communication, Control and Information Sciences (ICCISc)*. Vol. 1. IEEE, 2021.

Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems* 28 (2015).

Robinson, John B., and Deborah Herbert. "Integrating climate change and sustainable development." *International Journal of Global Environmental Issues* 1.2 (2001): 130-149.