Independent Study Report

Effects of Light on Microgreen Growth

Aleem Uddin
04/15/2024

Abstract

This study explores the optimization of light conditions for microgreen growth in vertical farming systems using machine learning and advanced image analysis techniques. The research primarily focuses on the impact of various light spectra and intensities on the growth rate of microgreens, leveraging artificial intelligence to enhance predictive accuracy and operational efficiency in indoor farming environments.

Utilizing a robust dataset comprising different light settings and corresponding plant responses, we developed and refined predictive models, including Random Forest and Artificial Neural Networks, to ascertain the most beneficial light conditions for microgreens. The study also integrated sophisticated image processing methods to evaluate plant growth facilitating real-time adjustments and precise control over farming conditions.

Introduction

1. Background

   Vertical farming is a sustainable agricultural practice that optimizes plant growth through controlled environment agriculture (CEA) technology. It uses artificial lighting, temperature control, and nutrient management to maximize production in limited spaces. One of the critical factors influencing plant health and productivity in these systems is lighting. The correct light spectrum can enhance photosynthesis, growth rates, and nutritional content of plants, making light optimization a central focus for research and technological development in indoor farming.

2. Research Objectives

   This study aims to delve deeper into the complex interactions between light conditions and microgreen growth. The current research focuses on growth rate and energy efficiency in response to various lighting conditions. The objectives of this study are:

A. To assess the impact of different light spectra and intensities on the growth of microgreens using a combination of machine learning models and image analysis techniques.
B. To develop and refine predictive models that can accurately forecast optimal lighting conditions for various types of microgreens.

3. Scope of the Report

This report presents a comprehensive analysis of the conducted experiments, the methodologies employed, and the findings obtained. It includes detailed discussions on the setup of the experimental designs, the data collection processes, the image analysis algorithms, and the development of machine learning models. The implications of our findings are discussed in the context of their potential to improve indoor farming practices and contribute to the broader field of agricultural technology.

Experimental Setup and Data Collection

In our experimental setup, we utilized four farm trays to grow mustard seeds under controlled environmental conditions. The variables such as water, temperature, and humidity were kept constant throughout the study to ensure uniformity and focus the analysis solely on the impact of lighting conditions.

1. Seed Germination

Mustard seeds were planted in the trays and allowed to germinate for three days in a dark environment to simulate natural initial growth conditions.

2. Automated System Setup

Post germination, the trays were moved to a setup equipped with automated RGB lighting and a camera system. The lights were programmed to turn on for specific durations every hour, during which the camera captured images of the trays.

3. Data Collection Duration

This process continued over the mustard seeds' typical growth cycle of seven days, resulting in a dataset comprising daily image captures from day one to day seven for each tray.

		
Day 1	**Day 4**	**Day 7**

Data Processing and Analysis

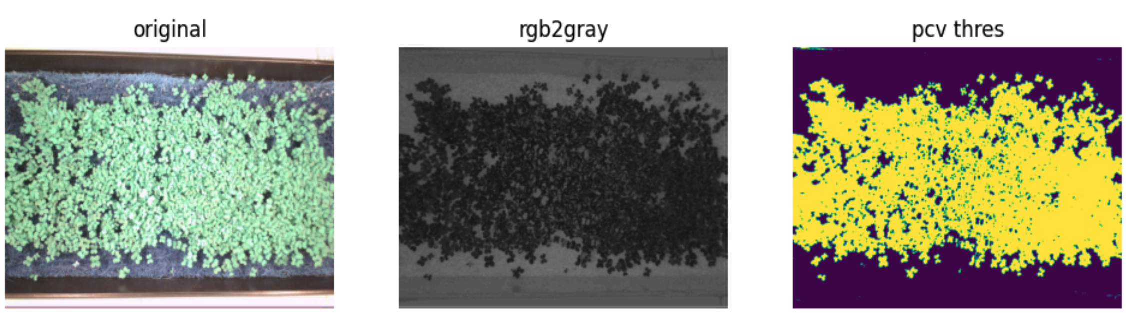
1. Image Processing:

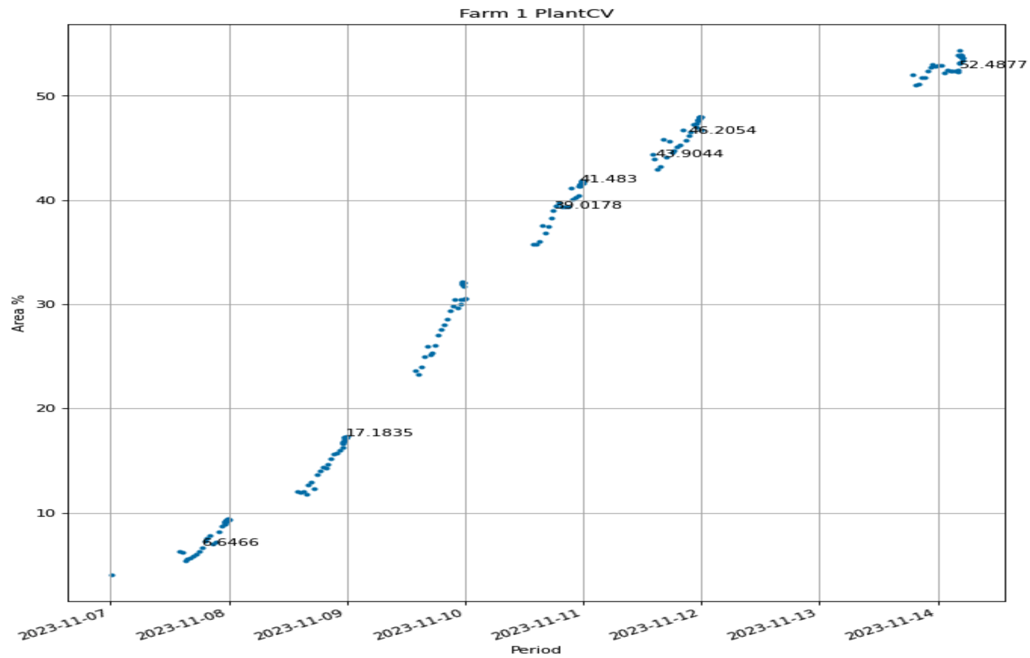
The image processing stage is crucial in extracting quantifiable data from the captured images of microgreen growth cycles. This stage involves several systematic steps to analyze the images, allowing us to determine the total plant area effectively. Our methodology leverages advanced image processing techniques to ensure accurate and reliable measurements.

We employed three primary methods for image segmentation, each chosen based on its ability to highlight plant areas effectively:

1.1. PlantCV Analysis

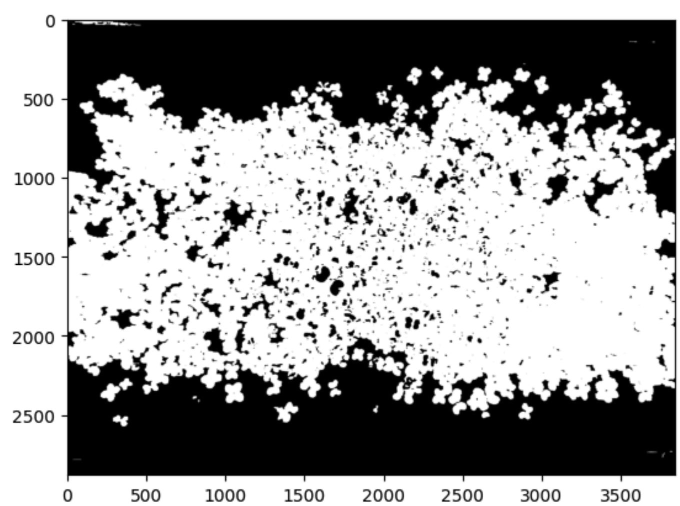
We utilized PlantCV to convert images to grayscale focusing on specific color channels that accentuate plant material. Following this, a binary threshold was applied to segment the plant from the background. Small objects and noise were removed using morphological operations to ensure a clean segmentation of plant area

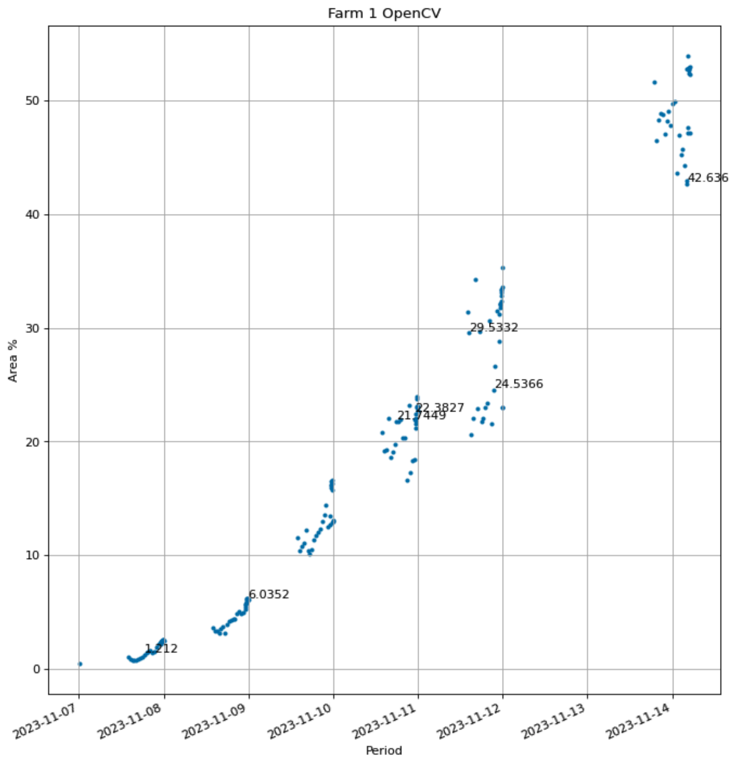




1.2. Open CV

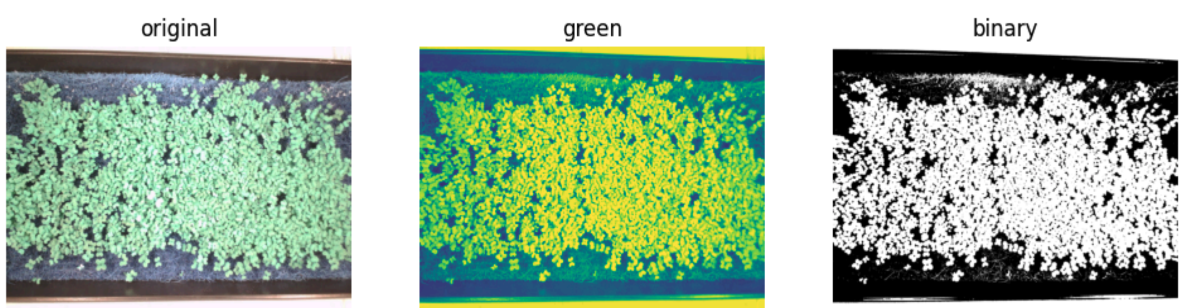
The OpenCV library facilitated the conversion of images to the HSV color space, which is particularly suited for processing plant images due to its sensitivity to color variations. We defined specific green color ranges to isolate plant matter and applied morphological operations to refine the segmentation. The contours of plant areas were then identified and used to calculate the total plant area.

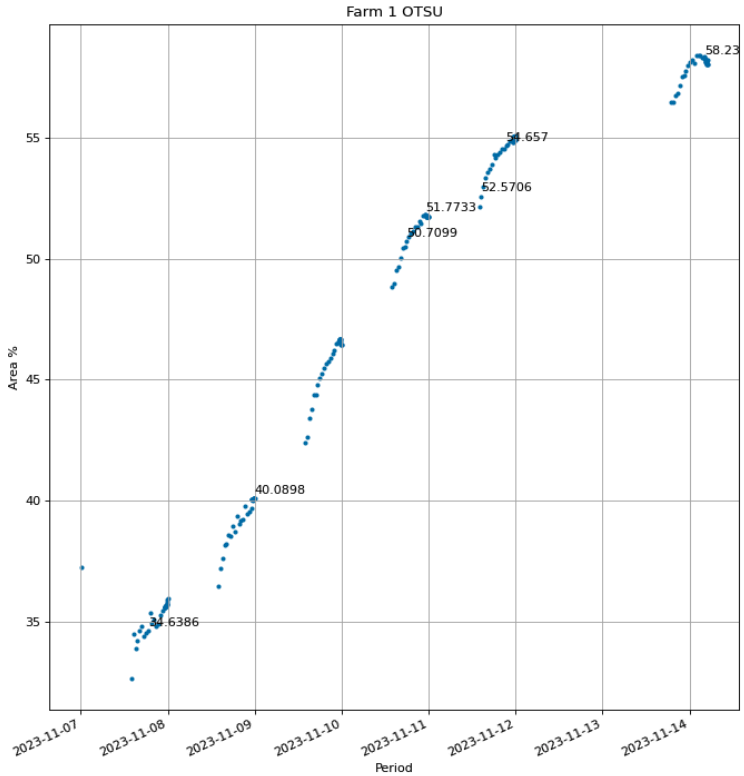




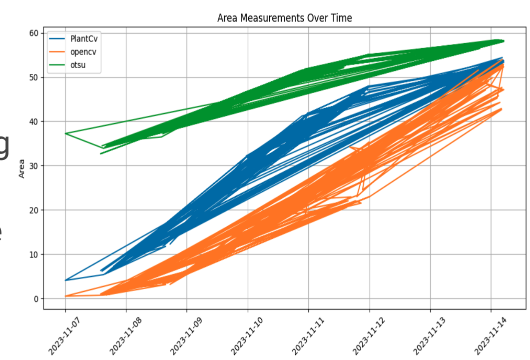
1.3. Otsu Thresholding

For a more automated approach, Otsu's thresholding was applied to the green channel of the images. This method automatically determines an optimal threshold, separating plant material from the background effectively.





Conclusion



The image processing methodology provided us with robust and precise measurements of plant growth areas, essential for the subsequent analysis stages. This data forms the foundation for our machine learning models, which predict growth outcomes based on various light conditions. The effectiveness of our image processing stage ensures that the inputs to these models are of high quality, thereby enhancing the reliability of our predictions.

2. CSV Data Compilation

After the image processing stage, where various methods were used to calculate plant area, the next crucial step involved compiling these results into a structured CSV file. This file serves as a dataset for subsequent machine learning analysis, containing comprehensive information about each image, the conditions under which it was captured, and the computed plant area.

2.1. Calculation of Average Plant Area

2.1.1. **Multiple Method Integration:** Plant area was calculated using three different image processing techniques to ensure accuracy and mitigate method-specific biases. These methods included PlantCV, OpenCV, and Otsu's thresholding.

2.1.2. **Average Area Calculation:** For each image, the areas calculated by each method were averaged to derive a final plant area value. This averaging approach helps smooth out anomalies that might occur if relying on a single method and provides a more reliable measure of actual plant growth.

2.2. CSV File Creation

Header Definition: The CSV file was structured with headers designed to capture not only the environmental conditions but also the results of the image analysis. The headers in the CSV file include:

Name: The identifier of the image or the experimental setup.

Date: The timestamp or date when the image was captured.

Light Conditions: Separate columns for different light spectra used during the experiment (R1, R2, R3 for different types of red light, B1, B2 for blue light, G for green, A for amber, and W for white light).

Area: The calculated average area of the plant material in the image.

DInMin: Duration in minutes from the start of the experiment or the specific light cycle.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Name	Date	R1	R2	R3	B1	B2	G	A	W	Area	DInMin
2	Farm 2	##########	0	102	0	0	165	0	0	0	324.98	6836
3	Farm 2	##########	0	102	0	0	165	0	0	0	332.47	6972
4	Farm 2	##########	0	102	0	0	165	0	0	0	402.74	11391
5	Farm 2	##########	0	102	0	0	165	0	0	0	372.6	7188
6	Farm 2	##########	0	102	0	0	165	0	0	0	466.59	11824
7	Farm 2	##########	0	102	0	0	165	0	0	0	439.61	11426
8	Farm 2	##########	0	102	0	0	165	0	0	0	414.76	5748
9	Farm 2	##########	0	102	0	0	165	0	0	0	462.82	11779
10	Farm 2	##########	0	102	0	0	165	0	0	0	210.84	2850
11	Farm 2	##########	0	102	0	0	165	0	0	0	292.81	3886
12	Farm 2	##########	0	102	0	0	165	0	0	0	340.1	8635
13	Farm 2	##########	0	102	0	0	165	0	0	0	319.64	4300
14	Farm 2	##########	0	102	0	0	165	0	0	0	365.57	8606
15	Farm 2	##########	0	102	0	0	165	0	0	0	395.11	5328
16	Farm 2	##########	0	102	0	0	165	0	0	0	264.15	3750
17	Farm 2	##########	0	102	0	0	165	0	0	0	312.26	8104
18	Farm 2	##########	0	102	0	0	165	0	0	0	328.55	6768
19	Farm 2	##########	0	102	0	0	165	0	0	0	330.96	4228
20	Farm 2	##########	0	102	0	0	165	0	0	0	289.71	3852
21	Farm 2	##########	0	102	0	0	165	0	0	0	330.64	4193
22	Farm 2	##########	0	102	0	0	165	0	0	0	370.04	11289
23	Farm 2	##########	0	102	0	0	165	0	0	0	373	5225
24	Farm 2	##########	0	102	0	0	165	0	0	0	365.55	8309
25	Farm 2	##########	0	102	0	0	165	0	0	0	390.55	8445
26	Farm 2	##########	0	102	0	0	165	0	0	0	355.91	7143
27	Farm 2	##########	0	102	0	0	165	0	0	0	175.99	2480
28	Farm 2	##########	0	102	0	0	165	0	0	0	399.3	11323
29	Farm 2	##########	0	102	0	0	165	0	0	0	211.02	2879
30	Farm 2	##########	0	102	0	0	165	0	0	0	200.75	2719
31	Farm 2	##########	0	102	0	0	165	0	0	0	320.02	4279

Results and Observations

1. **Machine Learning Analysis:**

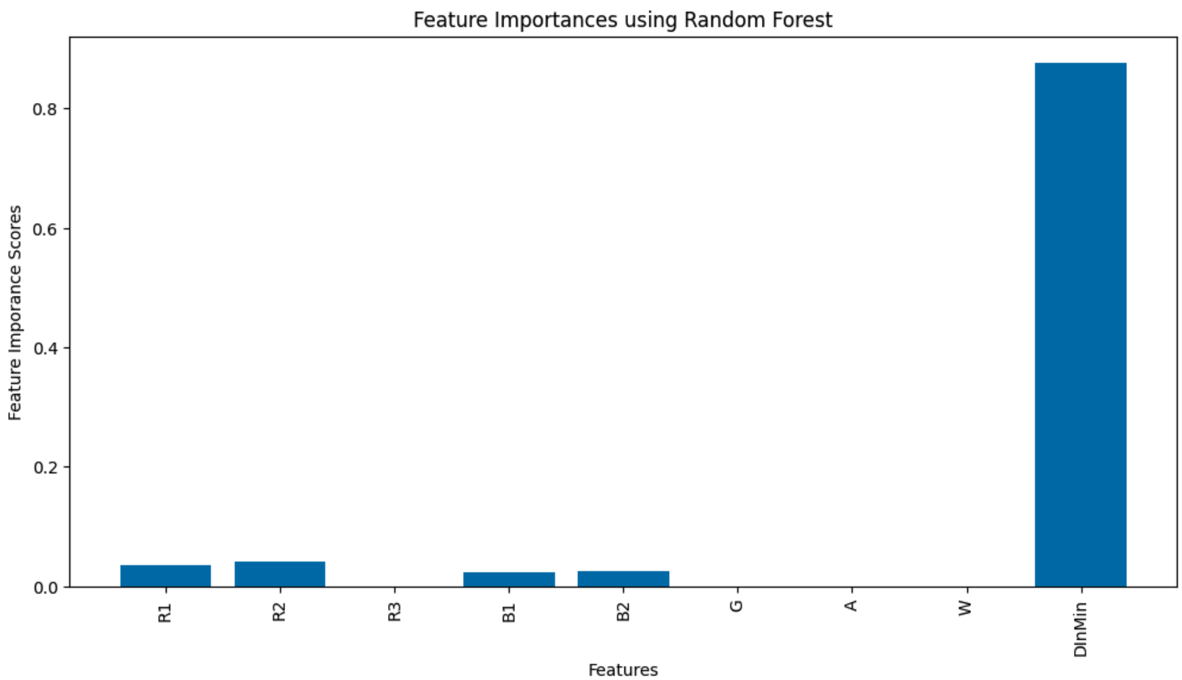
   Utilizing the processed data, we applied several machine learning models, including Linear Regression, PLSR, Lasso Regression, Ridge Regression, Random Forest, and Decision Tree, to analyze the data.

2. **Model Performance:**

   Due to the limited data set, only encompassing four distinct lighting conditions, most models did not perform optimally. The Random Forest and Decision Tree models showed the best results.

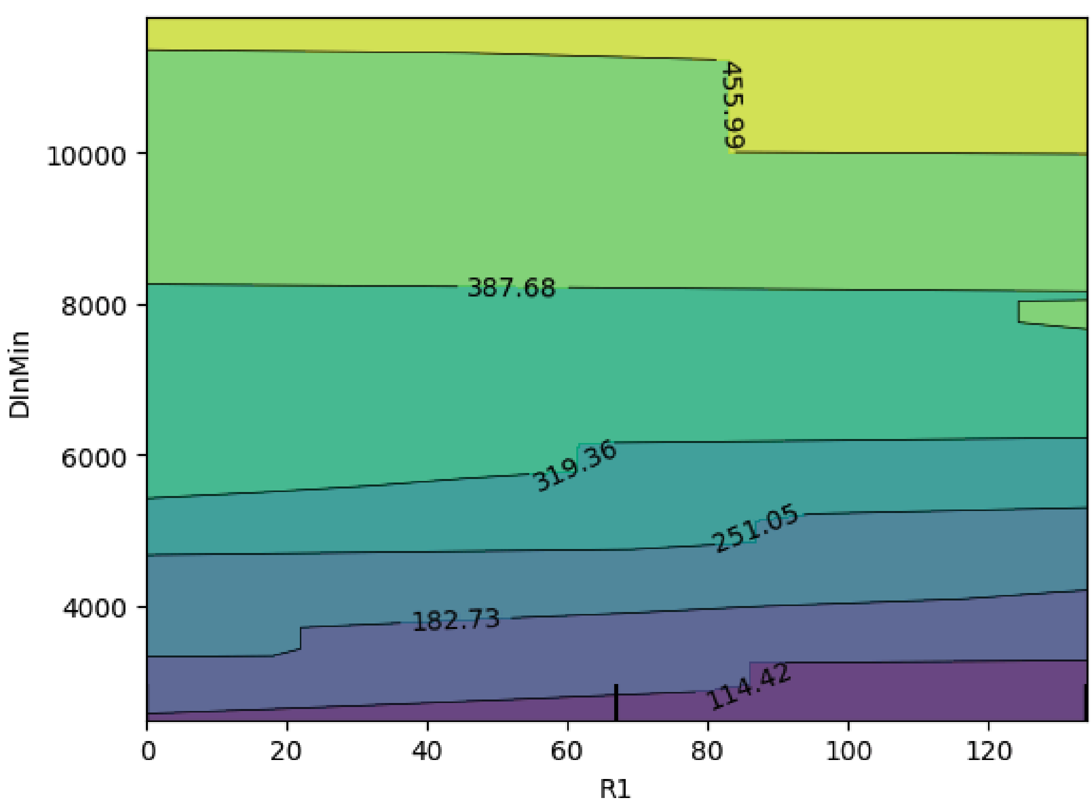
3. **Feature Importance:**

   Analysis of feature importance indicated that the models were more influenced by the 'Time' feature than by the varying light conditions, suggesting temporal factors played a more significant role than anticipated.

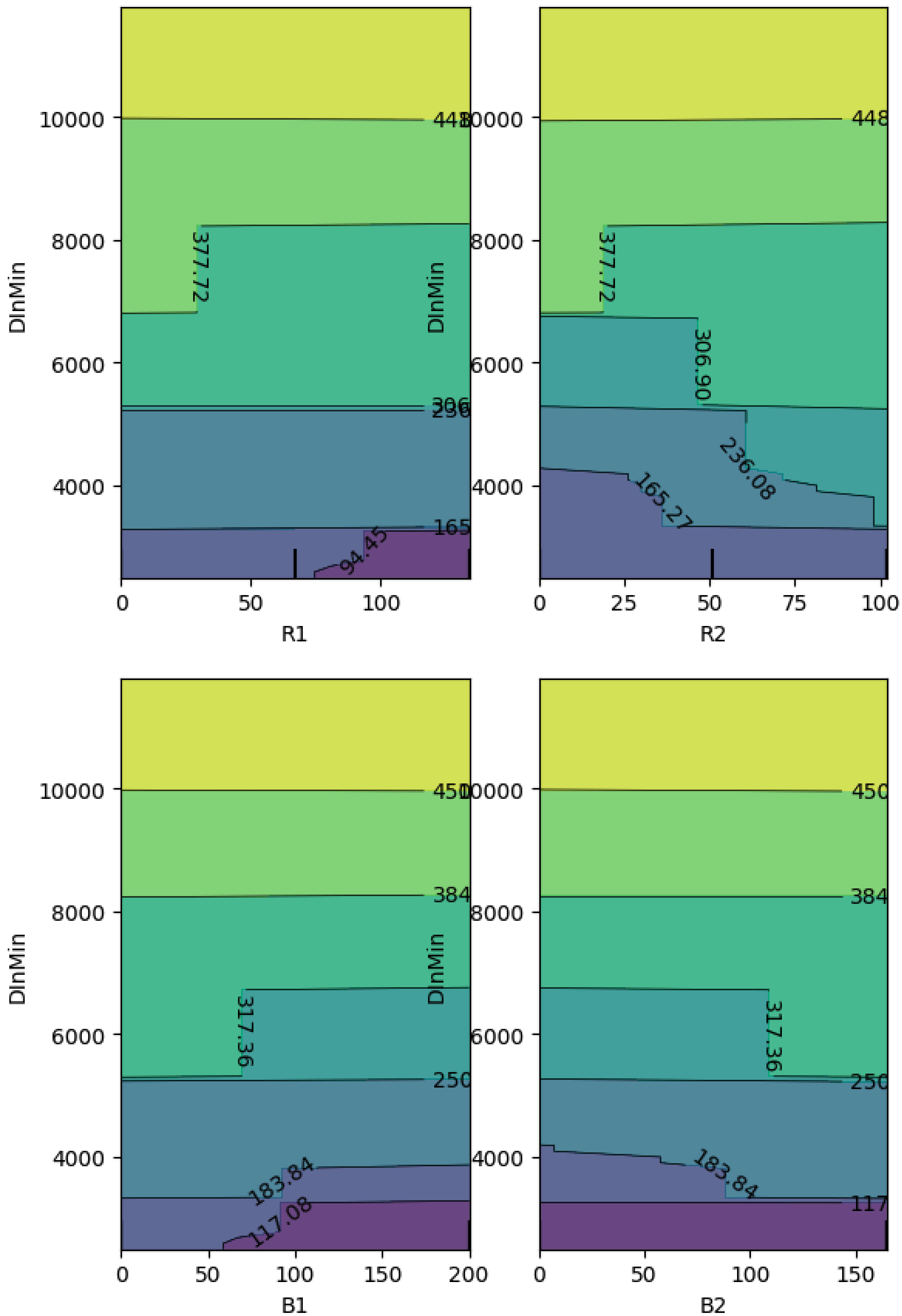


Feature Importances using Random Forest
Feature Imporance Scores
0.0
0.2
0.4
0.6
0.8
R1
R2
R3
B1
B2
G
A
W
DInMin
Features

4. PDP Plots:



455.99
387.68
319.36
251.05
182.73
114.42
10000
8000
6000
4000
DInMin
0
20
40
60
80
100
120
R1



DInMin
10000
8000
6000
4000
0
50
100
R1
448
377.72
306
236
165
94.45
25
75
R2
306.90
236.08
165.27
B1
450
384
317.36
250
183.84
117.08
150
200
B2
117

Future Study Directions:

To enhance the predictive accuracy and utility of the models, future experiments will explore a broader range of light conditions. The objective is to develop a robust model capable of predicting optimal growth conditions based solely on input light parameters.

Advantages

1. **Innovation**:

   Unlike other research, our study integrates both controlled lighting conditions and machine learning to analyze microgreen growth.

2. **Efficiency:**

   The methodology allows for straightforward calculation of plant area from images, offering a more efficient alternative to traditional manual measurements.

3. **Scalability**:

   This approach can be adapted for different types of plants, making it versatile for broader agricultural applications.

Limitations

1. **Data Limitations:**

   The main challenge was the small dataset, which limited the training capacity and generalizability of the machine learning models.

2. **Plant Diversity:**

   The study was confined to mustard plants, which may not represent the response of other microgreens or crops under similar conditions.

Conclusion

This study demonstrates a promising intersection of technology and plant science, using automated systems and machine learning to optimize microgreen growth conditions. With further research and a more extensive dataset covering various plant types and more diverse lighting conditions, this approach has the potential to significantly advance precision agriculture practices.