# ENVIRONMENTAL IMPACT ANALYSIS ON MONEY PLANT GROWTH USING IMAGE-BASED CLASSIFICATION

**BY**

**SANZIDA AKTER MUKTI**
**ID: 242-25-056**

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Masters of Science in Computer Science and Engineering

Supervised By

**Dr. Fizar Ahmed**
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised By

**Dr. Abdus Sattar**
Associate Professor & Director, M.Sc.
Department of CSE
Daffodil International University

**DAFFODIL INTERNATIONAL UNIVERSITY**

**DHAKA, BANGLADESH**

**SEPTEMBER 2025**

## APPROVAL

This Thesis titled **"ENVIRONMENTAL IMPACT ANALYSIS ON MONEY PLANT GROWTH USING IMAGE-BASED CLASSIFICATION"**, submitted by Sanzida Akter Mukti, ID No: 242-25-056 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 13-09-2025.

**BOARD OF EXAMINERS**

**Chairman**

**Dr. Sheak Rashed Haider Noori**
**Professor and Head**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

**Internal Examiner**

**Dr. Md. Zahid Hasan**
**Associate Professor**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

**Internal Examiner**

**Dr. Naznin Sultana**
**Associate Professor**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

**External Examiner**

**Mr. Nazibur Rahman**
**Head of IT Infrastructure**
Networld Bangladesh PLC

# DECLARATION

I hereby declare that, this project has been done by me under the supervision of **Dr. Fizar Ahmed, Associate Professor, Department of CSE, Daffodil** International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

**Supervised by:**

**Dr. Fizar Ahmed**
Associate Professor
Department of CSE
Daffodil International University

**Co-Supervised by:**

**Dr. Abdus Sattar**
Associate Professor & Director, M.Sc.
Department of CSE
Daffodil International University

**Submitted by:**

**Sanzida Akter Mukti**
ID: 242-25-056
Department of CSE
Daffodil International University

# ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes it possible to complete the final year thesis successfully.

I am really grateful and wish my profound indebtedness to **Dr. Fizar Ahmed, Associate Professor**, Department of CSE Daffodil International University, Dhaka, Deep Knowledge & keen interest of his supervisor in the field of "*Computer Vision*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express my heartiest gratitude to **Dr. Sheak Rashed Haider Noori**, **Head of the** Department of CSE, for his kind assistance in completing our project and also to other faculty member and the staff of CSE department of Daffodil International University.

I would like to thanks my entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

# ABSTRACT

The growth of indoor plants depends heavily on environmental conditions yet researchers have not conducted enough quantitative studies that combine visual data with contextual information. This research evaluates Money Plant (*Epipremnum aureum*) growth performance through the development of a time-series dataset We obtained daily plant photographs under both indoor and outdoor settings together with the measurements of Light Intensity, Temperature, Humidity and Water pH. The plant area measurements for growth assessment were obtained through image preprocessing steps that included contrast enhancement and HSV-based green masking and pot region exclusion. This study employed eight feature importance techniques to evaluate each parameter's contribution. This analysis revealed that Light Intensity stood as the primary growth factor while Water pH ranked as the secondary influential factors.

This research proves that using environmental features with image-derived growth metrics produces both interpretable and accurate predictions for plant growth. The established framework serves as a base for smart horticulture systems which optimize resource utilization and promote sustainable indoor plant maintenance.

Keywords: Indoor plant growth, Money Plant (Epipremnum aureum), Image-based analysis, Machine learning, Feature importance, Environmental factors, Smart horticulture.

# TABLE OF CONTENTS

| CONTENTS | PAGE |
|---|---|

# LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1
# Introduction

## 1.1 Introduction

Indoor plant cultivation has become increasingly popular in modern urban environments due to its combined aesthetic, ecological, and psychological benefits [25]. Indoor greenery has been shown to improve air quality, regulate humidity, and reduce stress, while also contributing to human productivity and well-being [1,2]. Among the wide variety of houseplants, the Money Plant (Epipremnum aureum) is one of the most widely adopted species, owing to its adaptability to diverse conditions and low maintenance requirements. Beyond its ornamental role, the Money Plant is also effective in improving indoor air quality by filtering pollutants, making it both a decorative and functional addition to modern living spaces [3].

Plant growth depends on several environmental parameters, such as light intensity, water availability, temperature, humidity, and soil quality [4,5]. While most plants are soil-based, the Money Plant in this study was cultivated hydroponically in water, making Water pH a critical growth factor instead of soil quality, the pH level is especially important because it affects the plant's ability to absorb nutrients [16]. Water-grown plants are more sensitive to environmental changes than those grown in soil. Traditional methods of plant care often rely on observation and can be inconsistent, which may not always yield the best results. Using a data-driven approach to measure how different conditions affect Money Plant growth can help improve plant care in controlled settings.

Recent advances in computational methods have enabled data-driven approaches to plant phenotyping, where image-based analysis plays a central role in measuring traits such as leaf area, canopy coverage, and biomass [6-8]. Preprocessing techniques like noise reduction, color segmentation, and morphological analysis allow the automated extraction of growth indicators from plant images [8]. At the same time, IoT and sensor-based monitoring provide accurate tracking of environmental variables, enabling a more comprehensive understanding of their influence on growth outcomes [9].

Despite these advances, most existing research has been concentrated on large-scale agriculture and crop production, with applications ranging from yield prediction to disease and stress detection [10-12]. In contrast, indoor ornamental plants have received limited attention, with studies often focusing narrowly on single factors such as light spectrum optimization [13] or irrigation control [14,15]. Moreover, many approaches fail to combine image-based growth measurements with environmental data, leaving a gap in systematic analysis for indoor horticulture.

In this study, A novel dataset was assembled, consisting of high-resolution time-series images of Money Plants captured daily, along with manual recordings of four key environmental parameters. Advanced image preprocessing techniques were applied, then the plant area was extracted from each image as a morphological growth indicator. This was used to derive a growth classification label, based on whether the plant's measured area exceeded a threshold representing the maximum average growth observed between indoor and outdoor conditions. Several Feature Importance techniques were applied then. To reduce bias from individual methods, a consensus voting mechanism ranked features by how often they were identified as most influential. This study helps improve indoor gardening by providing useful information for sustainable urban farming and smart plant care. By sharing a clear and repeatable method for tracking plant growth, this research bridges the gap between traditional gardening and modern, data-driven approaches.

## 1.2 Motivation

This research is motivated by the growing need for sustainable and efficient methods to cultivate indoor plants in urban areas [24]. Unlike earlier research [4], this study examines how various environmental factors interact. This approach offers practical tools for smart indoor gardening and new ideas for plant research. The main reasons for this work are:

- Scientific Gap: Most previous studies on indoor plant growth have used subjective observations or focused on just one factor at a time. As a result, there is still no systematic, data-driven way to assess how different environmental factors together affect plant performance.
- Emerging Technologies: New developments in computer vision and IoT now allow us to combine image-based plant analysis with environmental monitoring. This enables the creation of accurate and automated systems for predicting plant growth.
- Practical Applications: Urban residents and commercial horticulturists often struggle to keep indoor plants healthy for lack of proper guideline. A system that predicts important growth factors can provide valuable guidance for improved plant care.

- Environmental Sustainability: By highlighting the most critical factors such as light intensity and water pH, this framework supports resource-efficient plant care, reducing unnecessary energy and water use while promoting greener, healthier indoor spaces.

## 1.3 Research Objectives

a) To measure and compare the growth of Money Plants in indoor and outdoor environments.
b) To determine which environmental factor most strongly influences Money Plant growth.
c) To provide a data-driven basis for improving small-scale cultivation practices.
d) To enhance images by denoising, improving contrast, and removing non-plant artifacts for accurate area measurement.
e) To provide a foundation for automated, sustainable smart horticulture systems.

## 1.4 Research Questions

a) How do light intensity, temperature, humidity, and water pH affect Money Plant growth in indoor versus outdoor conditions?
b) Can plant area extracted from images serve as a reliable and automated growth indicator?
c) Which environmental factor emerges as the most crucial for plant growth when evaluated through multiple feature importance techniques?

## 1.5 Project Management and Finance

This research-based project didn't receive any financial support from any group or company.

## 1.6 Report Layout

Chapter 1 presents the research by describing its motivation, rationale, and main questions.
Chapter 2 reviews related works and highlights gaps in current literature.
Chapter 3 provides a detailed explanation of the experimental methodology.
Chapter 4 discusses the experimental results and examines the interpretability studies.
Chapter 5 shows impact on the environment and society along with the sustainability plan.
Chapter 6 provides a summary of the research and offers suggestions for future work.

# CHAPTER 2

# Background

## 2.1 Related works

Researchers have explored the uses of computer vision to study plant growth and optimize environmental conditions. Several studies demonstrate how integrating image-based analysis with environmental features can enhance plant monitoring systems. This study presents several studies from the literature that employed image datasets and environmental parameters to analyze and predict plant growth under varying conditions. Several studies have investigated the role of light intensity in plant growth, particularly in controlled settings like plant factories. For instance, Chen et al. [19] investigates how different light intensities (PPFD levels) affect lettuce photosynthesis and growth. Using ordinary least squares (OLS), the optimal constant light intensity was estimated at 140.8 $\mu mol \cdot m^{-2} \cdot s^{-1}$, which maintained photosynthesis without photoinhibition. A genetic algorithm support vector regression (GA-SVR) model was used to track how lettuce adapted to new light conditions, showing that short-term high light (290–480 $\mu mol \cdot m^{-2} \cdot s^{-1}$) improved adaptability in fluctuating environments. Growth experiments confirmed that 150–250 $\mu mol \cdot m^{-2} \cdot s^{-1}$ yielded the highest biomass. The study suggests low light levels are most energy-efficient for plant factories while higher short-term light boosts adaptability.

Similarly, Miao et al. [20] examined the effects of LED light intensities (300, 240, 180, and 120 $\mu mol \cdot m^{-2} \cdot s^{-1}$) on two lettuce cultivars (Crunchy and Deangelia) and one spinach cultivar (Shawen). They measured morphological, photosynthetic, and nutritional parameters, comparing results against natural light greenhouse conditions. While higher light intensities generally improved growth and nutrient content, they also induced stress symptoms like tipburn in lettuce and leaf curling in spinach. The study emphasized cultivar-specific requirements, identifying optimal light intensities of 240 $\mu mol \cdot m^{-2} \cdot s^{-1}$ for Crunchy and 180 $\mu mol \cdot m^{-2} \cdot s^{-1}$ for Deangelia, while Shawen performed poorly across all conditions.

Another study by Ncise et al. [14] investigates how light intensity, watering intervals affect the plant growth, nutrient content, and antifungal activity of hydroponically grown *Tulbaghia violacea*. Researchers cultivated plants under two light levels (0% and 40% shade) and three watering intervals (5, 14, and 21 days), measuring growth parameters, tissue nutrients, and extract efficacy against *Fusarium oxysporum*. They found that shorter watering intervals improved growth and nutrient uptake, while extended watering under low light enhanced antifungal activity. The study demonstrates that shading alleviates water-deficit stress and highlights the effects of light and water on medicinal plant quality.

Finally, Dai et al. [21] aimed to identify the optimal LED light intensity for balancing growth and nutritional quality of lettuce in a plant factory. Researchers grew lettuce under three light intensities (180, 210, and 240 $\mu mol \cdot m^{-2} \cdot s^{-1}$) and measured phenotypic traits, nutritional content, and photosynthetic parameters. They used radar charts, heatmaps, and a membership function for a comprehensive evaluation. The results showed that 210 $\mu mol \cdot m^{-2} \cdot s^{-1}$ (L2) promoted the best overall performance, maximizing yield and key nutrients, such as sugar and vitamin C. The study concluded that L2 is the optimal intensity for efficient, high-quality lettuce production in controlled environments.

Image-based phenotyping has become a critical approach for quantifying plant growth and morphology. Madhavi et al. [18], use a digital image preprocessing technique to measure the overlapping leaf area of ice plants. Researchers used HSV color space and threshold segmentation to calculate canopy area from images and then compared the results with the destructively measured total leaf area using a curvimeter. The method showed less than 10% overlapping leaf area and a strong correlation ($R^2 = 0.99$) between canopy and total leaf area, providing a reliable non-destructive approach for estimating leaf overlap in controlled environments.

Tu et al. [17] introduce a 2D in situ methodology for accurately measuring area of plant leaf using digital images. To improve precision, the method applies visual correction techniques and background color balancing, minimizing distortions and noise from varying light or imaging conditions. Experimental validation shows that the technique provides highly reliable leaf area measurements compared to traditional approaches. This method offers a low-cost, efficient solution for plant growth monitoring and phenotyping in agricultural research. The study didn't use a single overarching "model" but rather a pipeline of models and methods: a pinhole camera model with a distortion model for correction, and an HSV color model for image segmentation, all calibrated using Zhang's method.

Despite these advancements, significant gaps remain. Most studies focus on outdoor or greenhouse-grown crops, while ML and image-based phenotyping have been applied to stress detection and growth prediction, few studies use feature importance analysis to rank environmental factors. This limits the development of data-driven cultivation strategies tailored to indoor, hydroponic systems.

## 2.2 Research Gap

While several studies have investigated plant growth under different environmental conditions, most of the existing research has been conducted on large-scale food crops in highly controlled laboratory or greenhouse environments. Limited attention has been given to ornamental plants such as Money Plants, especially in real-world indoor and outdoor conditions. Moreover, previous works often rely on manual or invasive measurement of plant growth, which can be time-consuming and less accurate. Few studies have combined time-series environmental data with image-based plant area measurement to analyze growth trends. Additionally, the relative importance of environmental factors such as Light Intensity, Temperature, Humidity, and Water pH on plant growth remains underexplored in small-scale experimental setups. This study addresses these gaps by integrating environmental monitoring with image-based analysis to evaluate growth drivers in money plants.

## 2.3 Challenges

The main challenges encountered in this research are:

1. Dataset: The analysis utilized time-series images of Money Plant, accompanied by environmental data. This limits the applicability of the findings.
2. Image Processing: Accurate segmentation was challenging due to image noise, uneven lighting, and complex backgrounds. To address this, a strong preprocessing pipeline were developed.
3. Plant Area Measurement: Pixel-based area estimates were highly sensitive to segmentation errors. Reliable preprocessing was needed for accurate plant growth classification.
4. Feature Importance Analysis: Eight feature selection methods were employed to confirm the most significant factors influencing plant growth.

# CHAPTER 3

# Research Methodology

## 3.1 Working Process

This research follows a step-by-step process that combines visual and environmental data to thoroughly analyze the growth of the Money Plant (Epipremnum aureum). Figure 3.1 shows the complete workflow.

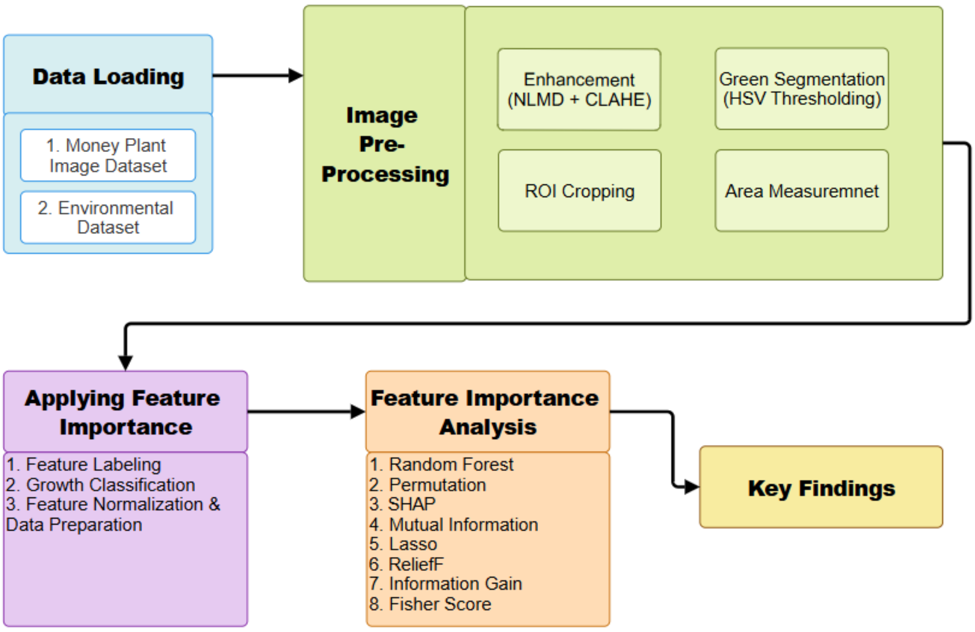


**Fig 3.1:** Research Workflow Diagram

The methodology begins with dataset creation, where high-resolution time-series images of the Money Plant were captured, evenly split between indoor and outdoor conditions. At the same time, an IoT device recorded three environmental factors: light intensity, temperature, and humidity. Water pH as well as the environment type (indoor or outdoor) were recorded manually. After collecting the images, the raw images underwent a multi-stage preprocessing pipeline consisting of image enhancement, green region segmentation via HSV thresholding, exclusion of the pot region, and final plant area measurement based on green pixel count. This area measurement served as the growth metric. Growth classification was then performed using a rule-based thresholding approach, labeling plants into High Growth or Low Growth categories. These morphological features were then combined with the environmental data from the IoT device to create a complete dataset. The dataset was then analyzed; eight feature importance techniques were used to systematically rank the contribution of environmental variables. Finally, a consensus voting framework was implemented to identify the most consistently influential features across all techniques. The experimental process concluded with visual interpretability analyses (consensus bar charts, correlation heatmaps). These results highlighted which environmental factors matter most for Money Plant growth indoors.

## 3.2 Dataset Description

The dataset is specially developed for this study is a unique time-series collection of high-resolution images of Money Plant. Images were divided equally between indoor and outdoor growth conditions. Sample images from Money Plant dataset shows in Figure 3.2.

Indoor Outdoor

**Fig 3.2:** Images from Money Plant Image Dataset.

For each image, five environmental factors were recorded: Light Intensity (lux), Temperature (°C), Humidity (%), Water pH, Environment Type (Indoor/Outdo1or). Light intensity, Temperature and humidity were measured continuously by the IoT device we made. Light Intensity was averaged over a 12-hour daily photoperiod, and the daily mean was calculated. Temperature and humidity were averaged over a 24-hour period, and water pH was measured daily using a pH meter. These data points were compiled into a CSV file where each row represented a single day with the associated image and environmental data. The CSV served as the input for analysis.

**Table 3.1:** Money plant Dataset

| **Image ID** | **Environment** | **Light Intensity** | **Temp (°C)** | **Humidity (%)** | **pH** |
|---|---|---|---|---|---|
| D-1.JPG | Indoor | 110 | 28 | 62 | 6.8 |
| D-30.JPG | Outdoor | 11628 | 28.1 | 84.25 | 5.92 |

Data Distribution of environmental parameters across the dataset is shown in Figure 3.3. Each subplot shows the variability of Light Intensity (lux), Temperature (°C), Humidity (%), and Water pH for indoor and outdoor plant growth conditions.

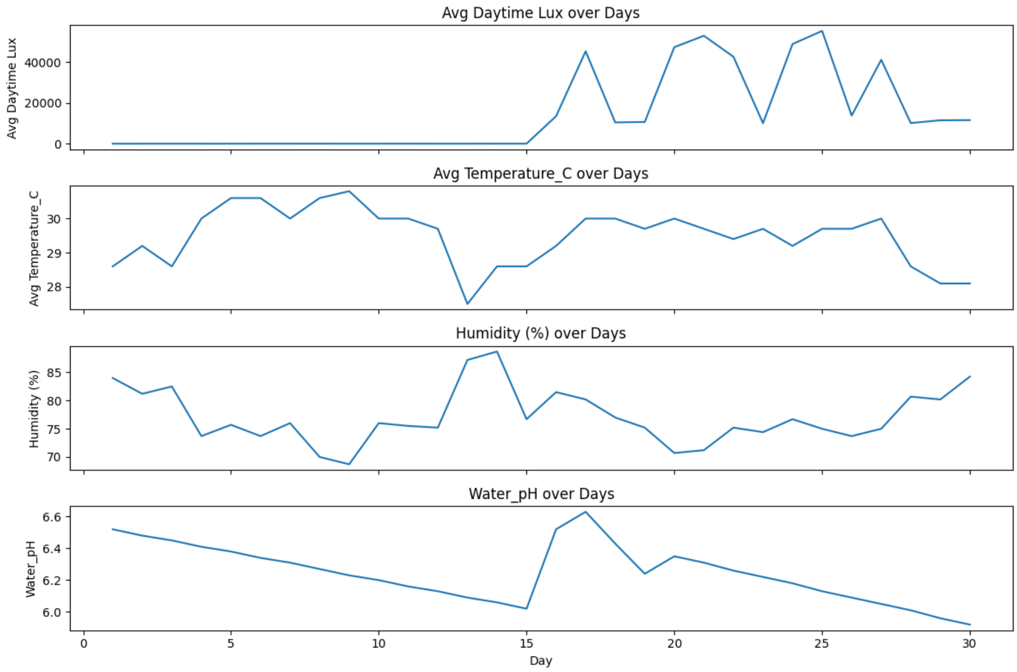


**Fig 3.3:** Data Distribution of Environmental Parameters

To measure the Light Intensity (lux), Temperature (°C), Humidity (%) the following IoT device has been made. It collects data and records that data every hour into the server.

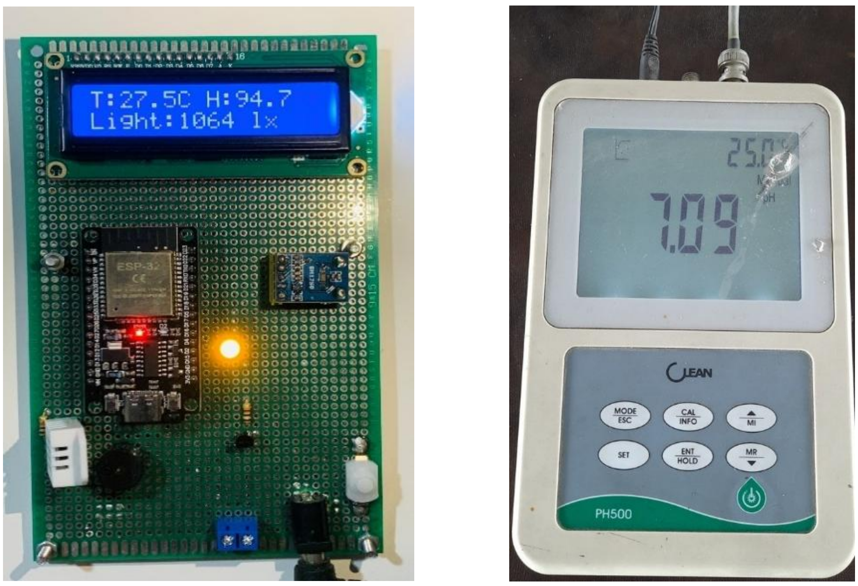


**Fig 3.4:** Images of IoT-based Environmental Monitoring Device and pH Meter

## 3.3 Image Pre-processing

Image pre-processing is essential for accurate and reliable plant area analysis using visual data. Raw plant images often include unwanted elements, such as pots, tables, inconsistent lighting, and noise, which can compromise segmentation quality. To solve these issues, a multi-step preprocessing pipeline was used. The process involved image enhancement, green region segmentation, removal of irrelevant areas, and measurement of plant area by counting pixels. Each step focused on isolating the green parts of the Money Plant, allowing for precise measurement and less noise in later classification. All steps involving Image preprocessing techniques are shown in Figure 3.5

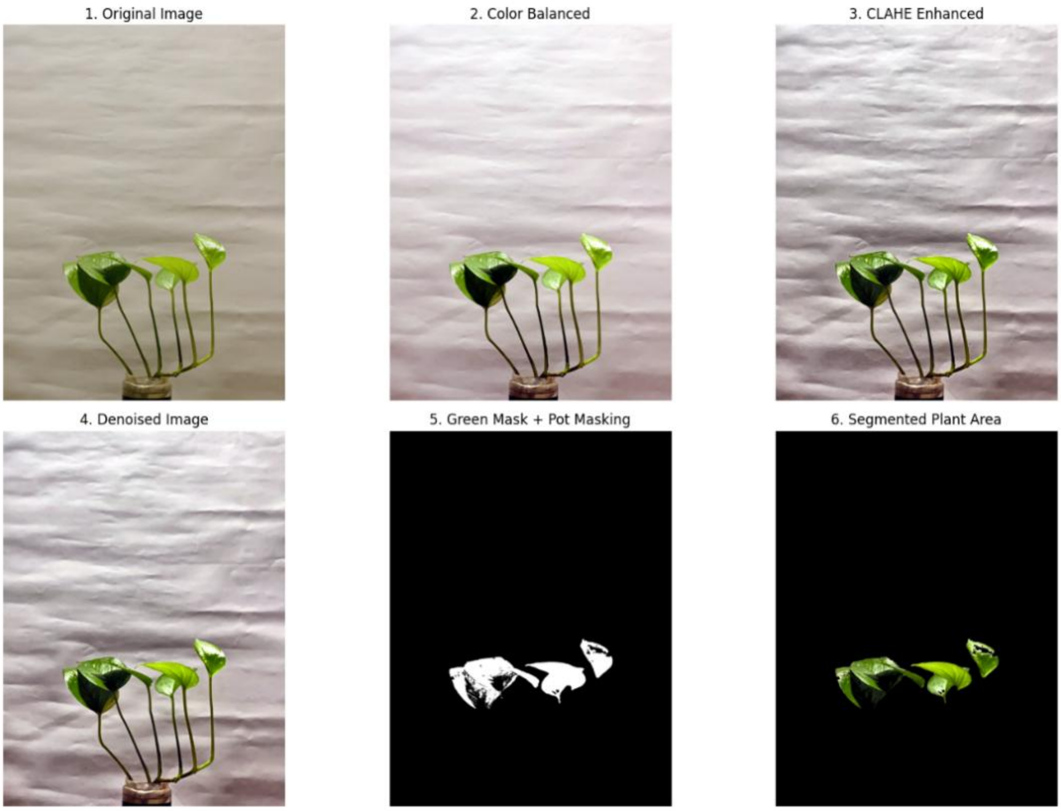


**Fig 3.5:** Image Preprocessing Pipeline

### 3.3.1 Image Quality Enhancement

Noise and low contrast often make it difficult to analyze images and identify it clearly [8]. To address this issue, morphological opening was used to reduce noise. Non-Local Means Denoising (NLMD) completely eliminate random noise while preserving the details of the plant leaves. Next, Contrast limited adaptive histogram equalization (CLAHE) enhanced the local contrast in plant regions, making green features more prominent. This method helped even low-light images keep important structure and color, which improved segmentation.

### 3.3.2 Green Region Masking and Segmentation

After morphological opening the images were then converted from RGB to HSV color space. A green mask was generated by applying HSV thresholds to isolate green hues. The threshold values were adjusted dynamically to handle different lighting conditions in the samples. This process separated the green foliage from the background, which was necessary for measuring plant area.

### 3.3.3 Pot Region Exclusion

Although green masking helped focus on the area of interest, portions of the pot often remained in the lower image section due to light reflections or overlapping plant parts. To fix this, bottom 10% was replaced of each mask with black, removing these unwanted areas. This way, only the plant leaves were included in the area estimation, and the image size stayed the same.

## 3.4 Plant Area Measurement

This study utilized image preprocessing techniques to measure plant growth by analyzing changes in green pixel area over time. The method involved generating binary green masks from plant images, where white pixels represented the plant's green regions. These pixel counts gave an objective numerical metric to track the growth progression. A custom function, compute_plant_area(), was implemented to process each image through multiple steps to isolate green plant pixels while excluding non-plant regions. The final pixel count, recorded as plant_area in a CSV file, provided a consistent method for measuring plant growth [17]. This method allowed us to monitor growth trends accurately and automatically by comparing daily changes in pixel area. It removed the need for visual observation and made our analysis more objective.

## 3.5 Feature Importance Techniques

To determine which environmental factors, have the greatest impact on Money Plant growth, we labeled, normalized, and selected features. This helped make sure the dataset was organized, balanced, and easy to analyze.

### 3.5.1 Feature Labeling (Growth Classification)

After calculating the plant areas for all images, we established a clear labeling method by determining the average plant area for both indoor and outdoor samples separately. We used the higher of these two averages as the threshold for classification. Samples with plant areas at or above this threshold were labeled as High Growth (1), and those below were labeled as Low Growth (0). This approach ensured that growth labels accurately reflected real differences in plant size, enabling models to distinguish between different environmental conditions. This binary target variable was then used for supervised classification tasks.

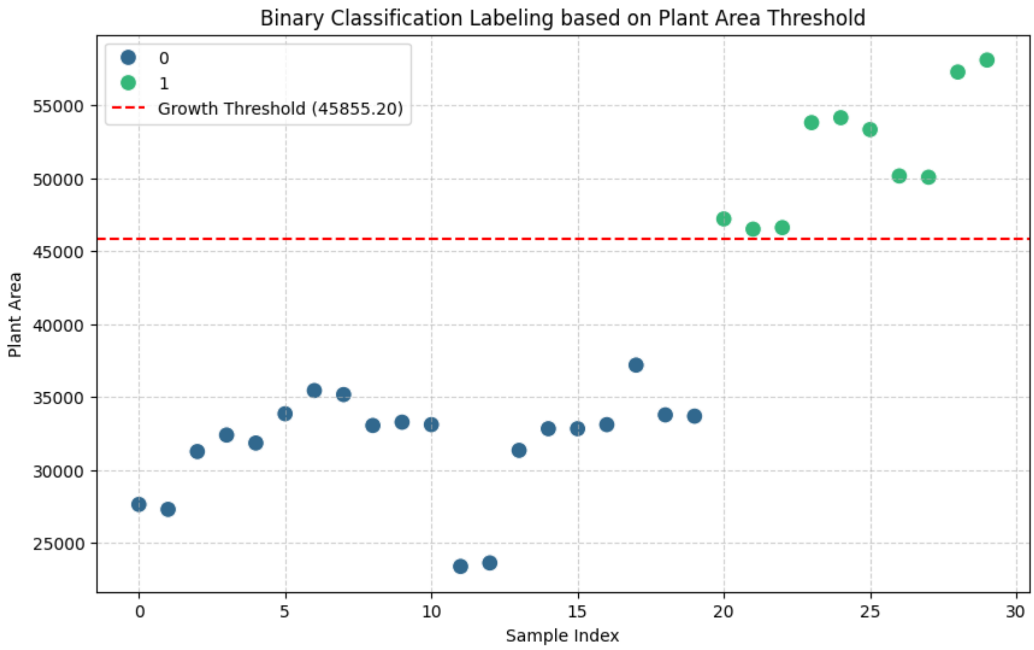


**Fig 3.6:** Binary Classification Labeling based on Plant Area Threshold

### 3.5.2 Feature Normalization & Data Preparation

To ensure that each environmental feature contributed equally, Light Intensity, Temperature, Humidity, and Water pH were normalized using the MinMaxScaler. This method scaled all features to values between 0 and 1, so that no single variable would outweigh the others in the model's learning process [22]. By normalizing the data, the model could learn more effectively, and it became easier to see which features were most important.

### 3.5.3 Feature Selection Techniques

Eight feature selection techniques were applied to determine the most important environmental features [23]. It helps us identified which environmental factors is the most crucial for Money Plant growth.

#### 3.5.3.1 Random Forest Importance

Random Forest is an ensemble learning method that constructs multiple decision trees and aggregates their predictions to improve accuracy and robustness. In this framework, feature importance is derived by measuring how much each feature contributes to reducing impurity (variance) when used for splitting nodes across the trees. Features that consistently produce better splits are assigned higher scores, making this method a reliable way to identify dominant variables in a dataset. Random Forest importance is particularly useful because it naturally handles non-linear relationships and interactions between features.

#### 3.5.3.2 Permutation Importance

Permutation importance is a model-agnostic technique that evaluates a trained model's reliance on each input feature. The method works by randomly shuffling the values of one feature at a time, thereby breaking its relationship with the target variable, and then measuring the decrease in model performance. A larger drop in accuracy or other evaluation metrics indicates that the feature is more important for predictions. This approach is valuable because it provides a direct, intuitive measure of feature relevance, independent of the model type.

#### 3.5.3.3 SHAP (Shapley Values)

SHAP (SHapley Additive exPlanations) is an advanced method based on cooperative game theory, designed to fairly distribute the contribution of each feature to a model's prediction. It calculates the marginal contribution of a feature by comparing predictions with and without that feature across all possible feature subsets. This ensures both global interpretabilities, by ranking features overall, and local interpretability, by explaining individual predictions. SHAP is widely regarded as one of the most rigorous approaches for feature importance, offering both transparency and consistency.

#### 3.5.3.4 Mutual Information (MI)

Mutual Information is a statistical measure that quantifies the dependency between two variables, capturing both linear and non-linear relationships. It evaluates how much knowing the value of one feature reduces uncertainty about the target variable. Unlike correlation, which captures only linear associations, mutual information can detect complex patterns, making it more versatile. Features with higher mutual information scores are considered more relevant for predicting the target.

#### 3.5.3.5 LassoCV

Lasso regression is a linear model that incorporates an L1 regularization penalty, which constrains the sum of the absolute values of the feature coefficients. This penalty has the effect of shrinking less important coefficients toward zero, effectively performing automatic feature selection during training. The LassoCV variant employs cross-validation to select the optimal regularization parameter, ensuring better generalization. As a result, only the most influential features remain with non-zero weights, simplifying the model while retaining predictive power.

#### 3.5.3.6 ReliefF Algorithm

The ReliefF algorithm identifies important features by iteratively sampling instances from the dataset and comparing them with their nearest neighbors. It evaluates how well each feature distinguishes between instances of the same class (nearest hits) and instances of different classes (nearest misses). Features that consistently differentiate between classes are assigned higher importance scores. ReliefF is particularly effective for detecting features that interact in subtle ways with others, making it suitable for complex datasets.

#### 3.5.3.7 Information Gain

Information Gain is an entropy-based measure that evaluates the usefulness of a feature by calculating how much it reduces uncertainty in the target variable when used for splitting the data. Commonly employed in decision tree algorithms, it guides the choice of attributes that best separate the data into homogeneous groups. A feature with higher information gain provides greater improvement in classification purity and helps the model make more accurate predictions. This makes it a fundamental and intuitive method for feature selection.

#### 3.5.3.8 Fisher Score

The Fisher Score is a statistical method that evaluates the discriminative power of features by comparing the ratio of inter-class variance to intra-class variance. Features that produce greater differences between classes while maintaining consistency within the same class achieve higher scores. This method is computationally simple yet highly effective for ranking features in classification problems. The Fisher Score is particularly useful when the goal is to identify features that clearly separate one class from another.

## 3.6 Experimental Methodology

In summary, the methodology integrates image-based growth measurement, threshold-based growth classification, and environmental factor analysis. The Plant area measurement went through a step-by-step image processing approach, which involved color balancing, CLAHE contrast enhancement, denoising, and HSV-based green segmentation to isolate leaf regions while excluding non-plant structures. The measured plant area was used as a growth indicator and to create binary growth labels (high or low growth). These labels, along with normalized environmental features including light intensity, humidity, temperature, and water pH, formed the supervised learning dataset for statistical interpretation. By combining multiple feature importance techniques with correlation analysis, the framework provided a reliable, interpretable, and data-driven way to determine the most critical environmental factors influencing Money Plant growth. Because these methods yielded different rankings, a voting system were employed to determine how often each feature was ranked as the most important.

### 3.7 Validation Strategy

To ensure the reliability of the results, a validation strategy was employed instead of a traditional ablation study, given the scope of this work. Since the methodology focuses on image-based classification and statistical analysis rather than complex multi-stage machine learning pipelines, the validation was performed through cross-checking results across multiple independent techniques. Specifically, eight feature importance methods were applied to evaluate environmental parameter influence. In addition, correlation analysis and statistical comparisons between indoor and outdoor growth conditions were conducted. By observing consistent trends across these diverse approaches, the study ensured robustness and reliability of the conclusions regarding the most influential environmental factors on Money Plant growth.

# CHAPTER 4

# Experimental Results and Discussion

## 4.1 Overview of Experimental Setup

The experiment was conducted in two different conditions: indoor (low-light) and outdoor (high-light). An IoT-based device was used to record environmental parameters including light intensity, temperature and humidity, and water pH was measured manually alongside daily images. These paired visual and environmental data served as the basis for classification and statistical analysis.

## 4.2 Evolution Methods

Growth evaluation was performed by calculating the average plant area for indoor and outdoor groups and using the higher mean value as the classification threshold. Images were then labeled as *High Growth* (≥ threshold) or *Low Growth* (< threshold). Feature importance techniques further highlighted which environmental variables contributed most significantly. Additionally, Statistical analysis, including mean, variance, and correlation heatmap among environmental factors, was carried out to determine their influence on growth and it validated the relationship between plant area and environmental parameters.

## 4.3 Results and Discussion

### 4.3.1 Results of Plant Segmentation and Area Calculation

The Plant area was successfully segmented using the image preprocessing pipeline. Visual inspection (Fig 4.1) showed accurate isolation of green foliage with minimal background interference. Each green pixel is counted to calculate the plant area. This method accurately excludes non-plant regions like pots and backgrounds.


**Figure 4.1:** Sample Image of Original Image, Preprocessed Image, and Area Extraction

After area extraction from the Money Plant images, the plant area was calculated for each image (Indoor and Outdoor). Sample of the plant area measurement is showing in Table 4.1.

**Table 4.1:** Calculated Sample Plant Area of Money Plant Images (Indoor vs Outdoor)

| Image ID | Environment | Plant Area |
|---|---|---|
| D-1.JPG | Indoor | 27622 |
| D-30.JPG | Outdoor | 58111 |

### 4.3.2 Results of Growth Classification

After area measurement, the average plant area (Mean) was calculated for each environment and found to differ significantly between the two environmental settings. These results suggest that Money Plants grow better in outdoor conditions.

**Table 4.2:** Plant Area Statistics: Mean, Min–Max Area Values for Indoor vs Outdoor Money Plant Samples

| Environment | Mean | min | max |
|---|---|---|---|
| **Indoor** | 31022.8 | 27622 | 33818 |
| **Outdoor** | 45855.2 | 33818 | 58111 |

To label growth, we used a threshold based on the higher mean value of 45,855 pixels. Images with plant areas at or above this value were labeled as ‘1’ for higher growth, while the rest were labeled as ‘0’ for lower growth.

**Table 4.3:** Sample Values of Plant Growth Measurement (Indoor vs Outdoor)

| Image ID | Environment | Plant Area | Growth Level |
|---|---|---|---|
| D-1.JPG | Indoor | 27622 | 0 |
| D-29.JPG | Outdoor | 58111 | 1 |

### 4.3.3 Results of Feature Importance Technique

Eight feature importance techniques were applied: Random Forest, permutation importance, SHAP, Mutual Information, Lasso, Relief, Information Gain, and Fisher Score. Across all four feature selection techniques, Light Intensity emerged as the most consistently important feature, being ranked first by 5 out of 8 methods. Water pH was ranked as the most important by 2 methods. Temperature and Humidity was less frequently prioritized across all methods. Table 4.4 shows Feature Importance Scores across all Model.

**Table 4.4:** Feature Importance Scores across All Model

| Feature | Light Intensity | Water pH | Temperature | Humidity |
|---|---|---|---|---|
| **Random Forest** | **1.00** | 0.3417 | 0.4999 | 0.00 |
| **Permutation** | 0.40 | **1.00** | 0.00 | 0.00 |
| **SHAP** | **1.00** | 0.4263 | 0.1181 | 0.00 |
| **Mutual Information** | **1.00** | 0.0874 | 0.4193 | 0.00 |
| **Lasso** | 0.00 | **1.00** | 0.1696 | 0.0241 |
| **ReliefF** | 0.7772 | **1.00** | 0.00 | 0.9090 |
| **Fisher Score** | **1.00** | 0.7761 | 0.00 | 0.0380 |
| **Information Gain** | **1.00** | 0.0874 | 0.4193 | 0.00 |

Consensus analysis confirmed that Light Intensity and Water pH were the dominant environmental variables influencing Money Plant growth under both conditions. This confirms the importance of light and placement environment in Money plant growth, supporting existing horticultural knowledge with data-driven validation. All methods showing relative influence of each feature visualized in bar plots Figure 4.2.

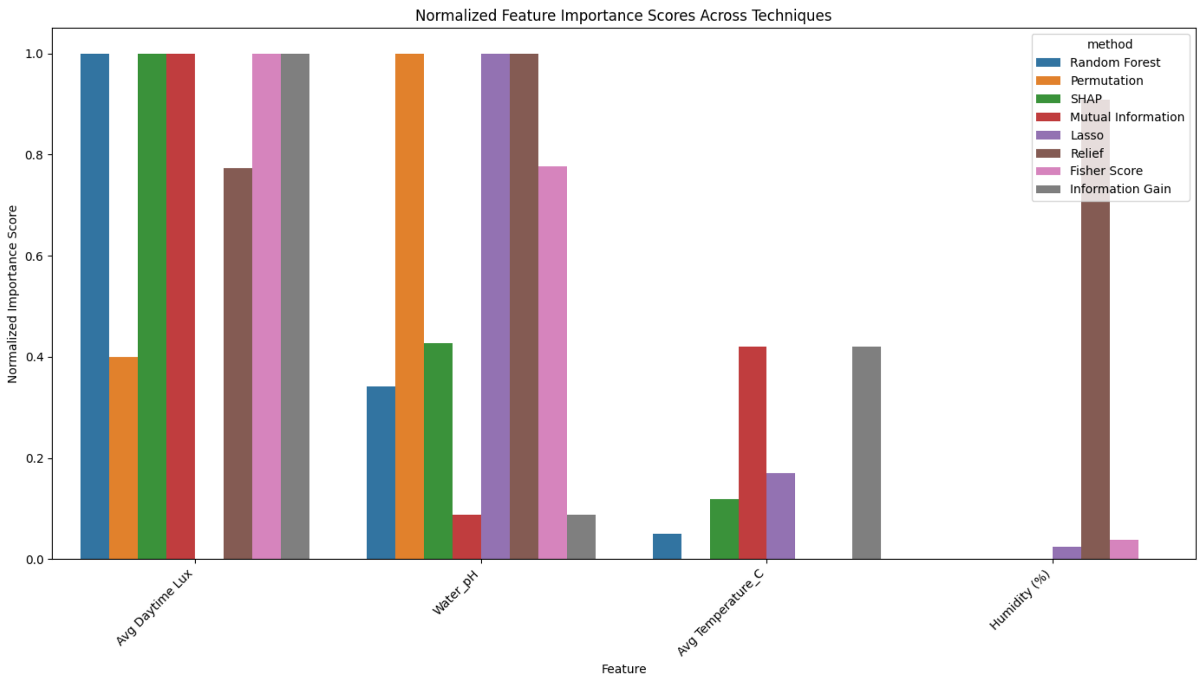


**Fig 4.2**: Comparison of Normalized Feature Importance Scores Across all Techniques

### 4.3.4 Statistical Analysis of Feature Selection Results

Correlation analysis was performed to examine the linear relationships between plant area and environmental variables. A heatmap was generated to visualize correlation strengths, while hypothesis testing was used to evaluate the statistical significance of each factor. This step validated the consistency of feature importance rankings with direct statistical evidence. The correlation heatmap showing the relationships between the input features. It shows that there is strong negative correlation among the Temperature and Humidity (-0.81). When temperature increases, Humidity decreases. The Temperature and Water pH has slightly positive correlation (0.38), that higher temperature is linked with slightly higher Water pH.

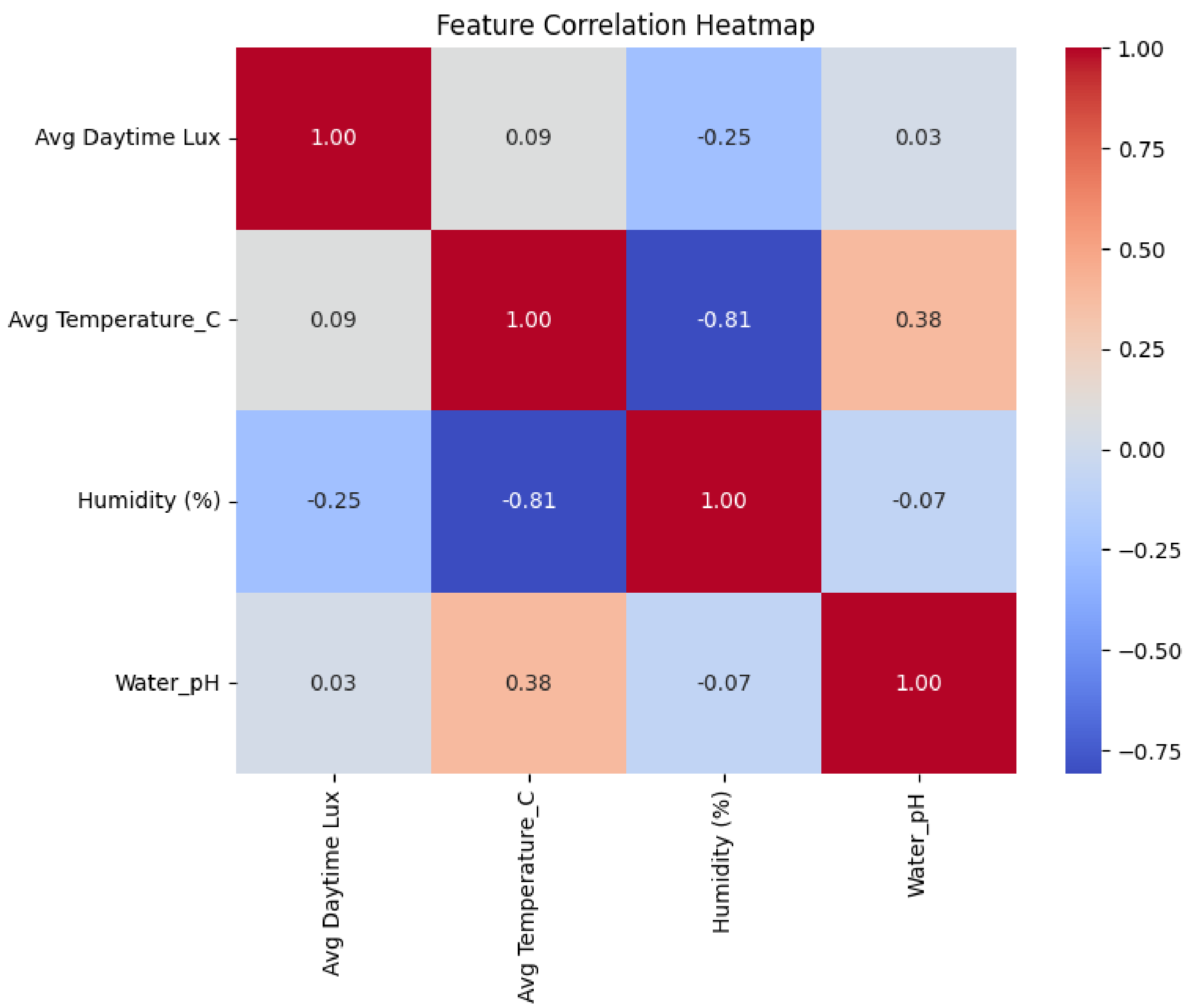


**Figure 4.3:** Confusion Matrix of Feature Correlation Heatmap

### 4.3.4 Visual Results and Interpretability

Visual analysis of sample plant images showed that the green-masking algorithm worked well for estimating plant area and tracking growth over time. The visualization confirmed that plants under outdoor conditions had visibly larger foliage and leaf spread.

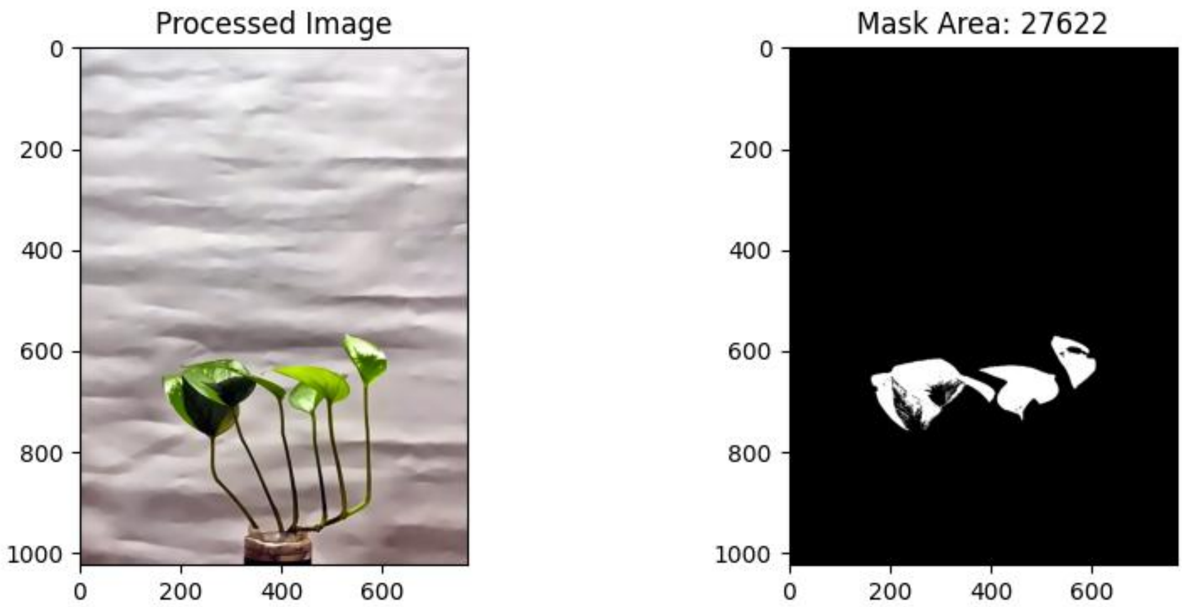


**Figure 4.3:** Sample Image of Money Plants Area Calculation (Indoor)

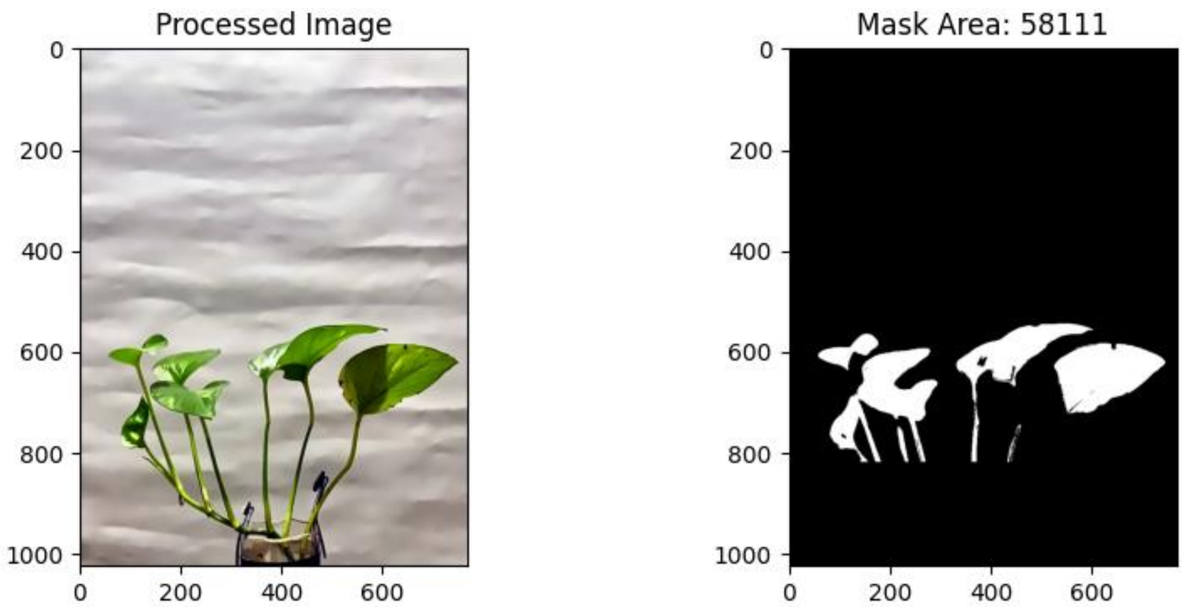


**Figure 4.3:** Sample Image of Money Plants Area Calculation (Outdoor)

### 4.3.5 Discussion

The experimental analysis yields important findings about the environmental factors that influence the growth of Money Plant. The image preprocessing and plant area measurement steps in the proposed pipeline produced reliable growth measurements. By using these area measurements to represent plant growth, the comparison between indoor and outdoor conditions showed clear differences. The analysis of feature importance revealed that light intensity was the most significant factor, ranking highest in several methods, including Random Forest, SHAP, Mutual Information, Fisher Score and Information Gain. This aligns with what is known about plant biology, where sufficient light is required for photosynthesis, especially for Money plants that are accustomed to low light levels. Water pH was the next most important factor, which makes sense given the plant's hydroponic environment and its need for the right nutrients in the water. These findings are consistent with existing research, which shows that light intensity is the main driver of photosynthesis and growth in money plants.

These results suggest that improving light and pH is key to better indoor Money Plant growth, while adjusting temperature and humidity can provide additional benefits. The threshold-based classification method facilitated the categorization of plant growth into 'higher' and 'lower' groups, and statistical tests demonstrated the reliability of this approach. In summary, combining plant area measurements from images with environmental monitoring is a valuable and effective method for assessing indoor Money Plant growth. The main finding that light intensity is the most important factor for growth adds to our understanding of plant science and provides practical advice for urban gardening. These results also highlight the value of creating data-driven and efficient systems for plant care, which can be applied to other types of ornamental plants and indoor farming.

# CHAPTER 5

## Impact on Society, Environment and Sustainability

### 5.1 Impact on Society

Indoor plants are now common in homes, offices, and public areas. People value them for their looks and for the positive effects they have on mental health and productivity. Research indicates that being surrounded by indoor plants can reduce stress, promote relaxation, and enhance focus at work. The Money Plant is especially popular because it is easy to care for and looks attractive. Still, when plants do not grow well because of poor conditions, people may give up on keeping them indoors.

This study directly addresses this challenge by providing a data-driven framework for evaluating the impact of environmental factors on Money Plant growth. By providing an evidence-based understanding of how variables such as light intensity, humidity, temperature, and water pH impact Money Plant growth, the study empowers plant owners to optimize conditions in a structured way rather than relying on guesswork. For households, this means improved plant survival rates and healthier, greener vegetation. On a larger scale, the research can inform the design of green urban infrastructure and indoor biophilic design practices, thereby supporting social well-being and sustainable city planning.

### 5.2 Impact on Environment

Combining image analysis with environmental monitoring helps support sustainability. This research demonstrates that it is possible to accurately track plant growth without requiring extensive resources or harming the plants. By matching plant care to what the plants actually need, the system helps save energy and water, which are often wasted in indoor gardening. In addition, successful indoor plant growth has broader ecological benefits. Indoor greenery helps regulate microclimates within buildings by improving humidity balance, and Healthy indoor plants also bring wider environmental benefits. They help control humidity and keep buildings cooler, which can lower the need for air conditioning. When more people grow healthy plants indoors, it can also support urban biodiversity and help clean the air inside buildings. The framework is practical for deployment on low-power devices or IoT-based monitoring systems, thereby ensuring scalability with minimal carbon footprint.

## 5.3 Sustainability Plan

This framework for monitoring Money plant growth using image-based classification prioritizes sustainability. To keep the system viable and eco-friendly over time, it focuses on several main areas:

a) Energy efficiency: The classification and feature analysis methods applied in this study (e.g., Random Forest importance, Lasso, Fisher Score) are computationally lightweight, ensuring reduced energy requirements.

b) Use of renewable energy: The system can be powered using renewable energy sources such as rooftop solar panels, making it suitable for eco-friendly urban households and indoor farms.

c) Resource Optimization: Automated monitoring of growth trends allows for precise control of lighting, irrigation, and ventilation, thereby minimizing unnecessary water and electricity usage.

d) Scalability with Low Environmental Footprint: The framework can be scaled to larger indoor horticultural setups or smart greenhouses without proportionally increasing energy consumption.

e) Long-Term Ecological Impact: By fostering healthier and more resilient indoor plant growth, the system supports urban greening efforts, reduces pollutant loads, and contributes to the development of sustainable urban ecosystems.

# CHAPTER 6

# Summary, Conclusion and Future Work

## 6.1 Summary of the Paper

This study examined the impact of various environmental conditions on the growth of Money Plants using an image-based classification method. Time-series images from both indoor and outdoor settings were collected, along with environmental records of Light Intensity, Temperature, Humidity, and Water pH. Then processed the images to reduce noise, improve contrast, and highlight green areas, which served as a measure of plant growth. Growth was categorized into two groups, high and low, based on a predetermined threshold derived from plant area measurements. Several methods, including Random Forest, SHAP values, Mutual Information, Lasso regression, ReliefF, Fisher Score, and Information Gain, were employed to determine which environmental factors were most significant. The results showed that Light Intensity had the biggest effect on Money Plant growth, followed by Water pH. Temperature and Humidity was less important. This supports the idea that light is crucial for plant health and also highlights the importance of maintaining good water quality in hydroponic setups.

## 6.2 Conclusion

This research demonstrates that combining images and environmental data is a reliable approach for studying plant growth. Unlike just observing plants, this method measures growth with clear data, helping people improve indoor plant care. Since light intensity was the most important factor, the study provides valuable insights for city dwellers, office planners, and plant experts. The framework is simple, non-invasive, and flexible, making it a good fit for city gardening, where resources may be limited. It could also be used for other types of indoor plants, both decorative and edible, and may help build more effective plant monitoring systems.

In summary, this study represents a crucial step in integrating computer vision and environmental monitoring to investigate indoor plant growth. The results benefit both researchers and practitioners by providing new knowledge and practical advice for cultivating plants indoors in a sustainable manner.

## 6.3 Limitations and Future Work

While the results are promising, the study has some limitations, including a small dataset and a reliance on manual measurement of water pH, which may limit the applicability of the findings. Future work should include a wider range of plant species, utilize real-time pH sensors, and conduct studies over longer periods to examine seasonal effects. As the dataset expands, machine learning and deep learning models can be applied to image-based phenotyping specifically for hydroponic ornamental plants, enhance predictive accuracy and automate growth classification, enabling more robust and scalable plant monitoring frameworks. Using advanced statistics and explainable AI can also make the results easier to understand and lead to more effective recommendations. Over time, this research could help create automated, efficient smart gardening systems that support sustainable city living.

**References**

[1] Bui, H., Park, J., Lee, E., Cho, W., Kwon, H., & Park, B. (2023). Assessment of the Air Cleaning Performance and Humidity and Temperature Control by Five Evergreen Woody Plants. *Atmosphere*. https://doi.org/10.3390/atmos14121819.

[2] Jiang, J., Irga, P., Coe, R., & Gibbons, P. (2024). Effects of indoor plants on CO2 concentration, indoor air temperature and relative humidity in office buildings. *PLOS ONE*, 19. https://doi.org/10.1371/journal.pone.0305956.

[3] Singh, K., Kumar, S., Paniteja, M., & Singh, S. (2019). Novel properties of Epipremnum aureum for treatment of fluoride-contaminated water. *SN Applied Sciences*, 1. https://doi.org/10.1007/s42452-019-0773-0.

[4] Ali, M., Yousef, A., Li, B., & Chen, F. (2021). Effect of Environmental Factors on Growth and Development of Fruits. *Tropical Plant Biology*, 14, 226 - 238. https://doi.org/10.1007/s12042-021-09291-6.

[5] Jia, J., Dai, Z., Li, F., & Liu, Y. (2016). How Will Global Environmental Changes Affect the Growth of Alien Plants? *Frontiers in Plant Science*, 7. https://doi.org/10.3389/fpls.2016.01623.

[6] Choudhury, S., Maturu, S., Samal, A., Stoerger, V., & Awada, T. (2020). Leveraging Image Analysis to Compute 3D Plant Phenotypes Based on Voxel-Grid Plant Reconstruction. *Frontiers in Plant Science*, 11. https://doi.org/10.3389/fpls.2020.521431.

[7] Choudhury, S., Maturu, S., Samal, A., Stoerger, V., & Awada, T. (2020). Leveraging Image Analysis to Compute 3D Plant Phenotypes Based on Voxel-Grid Plant Reconstruction. Frontiers in Plant Science, 11. https://doi.org/10.3389/fpls.2020.521431.

[8] Attari, H., & Ghafari-Beranghar, A. (2018). An Efficient Preprocessing Algorithm for Image-based Plant Phenotyping. https://doi.org/10.20944/preprints201804.0209.v1.

[9] Wu, Y., Wang, X., Wang, M., Liu, X., & Zhu, S. (2024). Time-Series Forecasting of PM2.5 and PM10 Concentrations Based on the Integration of Surveillance Images. *Sensors (Basel, Switzerland)*, 25. https://doi.org/10.3390/s25010095.

[10] Singh, A., Awasthi, A., Badola, U., Bidwe, R., & Mishra, S. (2025). A Novel Hybrid Approach to Crop Yield Prediction: Combining Deep Learning Efficiency with Statistical Precision. *International Journal of Computing and Digital Systems*. https://doi.org/10.12785/ijcds/1571032941.

[11] Huang, S., Liu, Q., Wu, Y., Chen, M., Yin, H., & Zhao, J. (2024). Edible Mushroom Greenhouse Environment Prediction Model Based on Attention CNN-LSTM. *Agronomy*. https://doi.org/10.3390/agronomy14030473.

[12] Elvanidi, A., & Katsoulas, N. (2022). Machine Learning-Based Crop Stress Detection in Greenhouses. *Plants*, 12. https://doi.org/10.3390/plants12010052.

[13] Pennisi, G., Pistillo, A., Orsini, F., Cellini, A., Spinelli, F., Nicola, S., Fernández, J., Crepaldi, A., Gianquinto, G., & Marcelis, L. (2020). Optimal light intensity for sustainable water and energy use in indoor cultivation of lettuce and basil under red and blue LEDs. *Scientia Horticulturae*, 272, 109508. https://doi.org/10.1016/j.scienta.2020.109508.

[14] Ncise, W., Daniels, C., & Nchu, F. (2020). Effects of light intensities and varying watering intervals on growth, tissue nutrient content and antifungal activity of hydroponic cultivated Tulbaghia violacea L. under greenhouse conditions. *Heliyon*, 6. https://doi.org/10.1016/j.heliyon.2020.e03906.

[15] Kittipornkul, P., Treesubsuntorn, C., Kobthong, S., Yingchutrakul, Y., Julpanwattana, P., & Thiravetyan, P. (2023). The potential of proline as a key metabolite to design real-time plant water deficit and low-light stress detector in ornamental plants. *Environmental science and pollution research international*. https://doi.org/10.1007/s11356-023-27990-3.

[16] Singh, K., Kumar, S., Paniteja, M., & Singh, S. (2019). Novel properties of Epipremnum aureum for treatment of fluoride-contaminated water. *SN Applied Sciences*, 1. https://doi.org/10.1007/s42452-019-0773-0.

[17] Tu, L., Peng, Q., Li, C., & Zhang, A. (2021). 2D In Situ Method for Measuring Plant Leaf Area with Camera Correction and Background Color Calibration. *Sci. Program.*, 2021, 6650099:1-6650099:11. https://doi.org/10.1155/2021/6650099.

[18] Madhavi, B., Bhujel, A., Kim, N., & Kim, H. (2022). Measurement of Overlapping Leaf Area of Ice Plants Using Digital Image Processing Technique. *Agriculture*. https://doi.org/10.3390/agriculture12091321.

[19] Chen, D., Zhang, J., Zhang, Z., Wan, X., & Hu, J. (2022). Analyzing the effect of light on lettuce Fv/Fm and growth by machine learning. *Scientia Horticulturae*.

[20] Miao, C., Yang, S., Xu, J., Wang, H., Zhang, Y., Cui, J., ... & Ding, X. (2023). Effects of light intensity on growth and quality of lettuce and spinach cultivars in a plant factory. *Plants*, *12*(18), 3337.

[21] Dai, M., Tan, X., Ye, Z., Ren, J., Chen, X., & Kong, D. (2024). Optimal light intensity for Lettuce growth, quality, and photosynthesis in plant factories. *Plants*, *13*(18), 2616.

[22] Shantal, M., Othman, Z., & Bakar, A. (2023). A Novel Approach for Data Feature Weighting Using Correlation Coefficients and Min-Max Normalization. *Symmetry*, 15, 2185. https://doi.org/10.3390/sym15122185.

[23] Huang, J., Qian, W., Vong, C., Ding, W., Shu, W., & Huang, Q. (2023). Multi-Label Feature Selection via Label Enhancement and Analytic Hierarchy Process. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7, 1377-1393. https://doi.org/10.1109/TETCI.2022.3231655.

[24] Ampim, P., Obeng, E., & Olvera-Gonzalez, E. (2022). Indoor Vegetable Production: An Alternative Approach to Increasing Cultivation. *Plants*, 11. https://doi.org/10.3390/plants11212843.

[25] Priya, U., & Senthil, R. (2024). Enhancing Sustainable Thermal Comfort of Tropical Urban Buildings with Indoor Plants. Buildings. https://doi.org/10.3390/buildings140823

242-25-056

ORIGINALITY REPORT

| 11% | 8% | 5% | 5% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Submitted to Daffodil International University<br>Student Paper | 2% |
| 2 | dspace.daffodilvarsity.edu.bd:8080<br>Internet Source | 2% |
| 3 | dokumen.pub<br>Internet Source | 1% |
| 4 | Submitted to University of Bradford<br>Student Paper | <1% |
| 5 | www.mdpi.com<br>Internet Source | <1% |
| 6 | econpapers.repec.org<br>Internet Source | <1% |
| 7 | Submitted to Liverpool John Moores University<br>Student Paper | <1% |
| 8 | Deema F. Turki, Ahmad F. Turki. "Machine Learning-Based Identification of Phonological Biomarkers for Speech Sound Disorders in Saudi Arabic-Speaking Children", Diagnostics, 2025<br>Publication | <1% |
| 9 | Elif Setenay Cutur, Neslihan Gokmen Inan. "Multi-class Classification of Retinal Eye Diseases from Ophthalmoscopy Images Using Transfer Learning-Based Vision | <1% |