PLNet: Light Recipe Design for Indoor Farming through Generative Deep Learning

Tanvi Verma, Ricardo Shirotta Filho
Institute of High Performance Computing (IHPC),
Agency for Science, Technology and Research (A*STAR), Singapore
{vermat, ricardosf}@ihpc.a-star.edu.sg

Abstract—Indoor farming has emerged as a promising solution for year-round cultivation and efficient resource utilization in crop production. Achieving optimal plant growth and quality in indoor environments requires precise control of light conditions. This study introduces PLNet, a generative deep learning method for the designing the light recipes specifically tailored for indoor farming.

Leveraging the power of deep neural networks, our approach establishes complex connections between light spectra and plant growth characteristics. Initially, a biomass estimator model is trained using a diverse dataset encompassing different light recipes and corresponding plant responses. Subsequently, a generative model is trained using the estimator model as a foundation, enabling the generation of optimal light spectra to achieve desired growth outcomes. This novel generative method offers an efficient and effective approach to formulating light recipes for indoor farming. By reducing the reliance on traditional trial-and-error methods, our method saves significant time and resources.

The presented generative deep learning method holds great potential for advancing the design of light recipes in indoor farming. Leveraging the capabilities of deep neural networks facilitates more targeted and efficient optimization of light conditions, resulting in improved crop yield and quality for a variety of leafy green crops. The findings of this study contribute to the ongoing efforts in enhancing productivity and sustainability in indoor cultivation practices.

Index Terms—Indoor farming, Generative deep learning, Light treatment, Light recipe design.

I. INTRODUCTION

Indoor farming has emerged as a transformative approach to agricultural practices, offering a viable solution for year-round crop cultivation and addressing challenges related to land scarcity and climate variability. By providing controlled environments, indoor farming enables precise manipulation of various growth parameters, including temperature, humidity, and most importantly, light conditions. Light plays a crucial role in plant growth and development, influencing photosynthesis, morphology, and nutritional content.

Traditionally, light recipes for indoor farming have been formulated through empirical approaches, relying on trial-and-error methods and expert knowledge. However, this approach has its limitations as it relies on a small set of predefined recipes and may overlook the vast space of possible light combinations. Furthermore, the complex and nonlinear relationships between light spectra and plant responses make it challenging to optimize light conditions effectively. This motivated us to leverage deep learning to enhance the design of light recipes in indoor farming.

Deep learning, a subfield of machine learning, has shown remarkable success in various domains, including computer vision, natural language processing, and speech recognition. Its ability to automatically learn hierarchical representations from large datasets and capture intricate patterns makes it a promising tool for tackling complex problems in agriculture. While deep learning has been extensively applied to image analysis and yield prediction in plant science [1]–[4], its utilization for the design of light recipes in indoor farming remains relatively unexplored.

In this study, we propose Plant Light Network, PLNet, a generative deep learning framework for the optimization of light recipes in indoor farming. Our approach aims to learn the underlying relationships between light spectra and plant growth characteristics, enabling the generation of tailored light recipes to achieve higher biomass yields. By leveraging the power of deep neural networks, we aim to overcome the limitations of traditional empirical approaches and provide a more efficient and effective solution for formulating light recipes in indoor farming.

The primary objective of this study is to demonstrate the efficacy of the generative deep learning method in designing optimized light recipes for indoor farming. Our methodology involves leveraging a plant growth dataset that captures a range of growth parameters, such as biomass and leaf area. These parameters are observed in saplings as they undergo development under various light recipes. We first use this dataset to train a biomass estimator model that estimates the plant growth based on given light conditions. Subsequently, the proposed generative model PLNet is trained using the biomass estimator model as a foundation, allowing the generation of light spectra that maximize the biomass yield of plant. In the training process of PLNet, we incorporate two distinct regularization terms. This serves two purposes: firstly, to expedite the convergence of the model, and secondly, to guarantee that the generated light recipe curves exhibit a smooth profile while maintaining the Photosynthetic Photon Flux Density (PPFD) within the desired range.

Through qualitative analysis we show that the generative deep learning method is a promising approach to design light recipes for indoor farming. Overall, this study represents a significant step towards leveraging deep learning for the
optimization of light recipes in indoor farming. By harnessing the power of deep neural networks, we aim to enhance productivity, resource efficiency, and sustainability in indoor cultivation, contributing to the ongoing efforts in transforming the future of agriculture.

II. LIGHT RECIPE

A light recipe refers to the specific combination of wavelengths and corresponding intensities of light that are used for plant growth and development in indoor farming settings. PAR (Photosynthetically Active Radiation) refers to the spectral range of solar radiation that is essential for photosynthesis in plants. This range generally covers the visible light spectrum. The intensity of light treatment is represented using PPFD (Photosynthetic Photon Flux Density) which is a measure of the amount of light energy in the PAR range that is available to plants for photosynthesis. PPFD is expressed in units of micromoles of photons per square meter per second ($\mu$mol/$m^2s$), and it indicates the number of photons that fall on a given area of plants per second. Figure 1 illustrates an instance of a light spectrum with a PPFD of approximately 145 $\mu$mol/$m^2s$. This value corresponds to the integral under the curve, symbolizing the total PPFD. The spectrum spans in the wavelength range of 380-780 nm encompassing the crucial colors necessary for photosynthesis.

III. RELATED WORK

Light optimization in indoor farming has been a subject of considerable research interest in recent years. Traditional approaches have relied on empirical methods, where growers and researchers manually adjust light spectra and intensities based on prior knowledge and trial-and-error experiments. Generally, the impact of light recipes on plant growth has been investigated through the selection of a limited number of manually designed recipes. Researchers have typically observed and compared various plant growth parameters to evaluate the effectiveness of these recipes. For instance, [5] utilized four distinct light recipes to examine the growth of choy sum and established that LED light intensity and spectrum both influenced growth, with the red-blue light treatment producing the highest shoot biomass at 160 $\mu$mol/$m^2s$. Meanwhile, [6] studied the effect of sole-sourced LED and mineral nutrient fertility treatment on Chinese kale and observed that plants showed superior accumulation of sulfur, boron, and zinc in the root tissue under the 10% blue/90% red LED light recipe, while iron concentrations were highest in the 40% blue/60% red LED light recipe. Another research on Chinese kale by [7] found that plants grown under fluorescent/incandescent light recipe had significantly higher shoot fresh and dry mass. [8] studied the effects of UV-A (ultraviolet-A radiation) irradiation on the cultivation and quality of microgreens and found that supplementing light recipes with UV-A irradiation resulted in increased leaf area and fresh weight of the plants. In a similar study, [9] explored the impact of supplementing red and blue light with UV-A on the growth of Kale. The findings indicated that the addition of UV-A positively influenced both the growth and quality of Kale. In a separate study on pak choi, [10] investigated the effect of light recipes on the regulation of carotenoid levels and discovered that blue, red, and white light had varying impacts on carotenoid composition. Although all of these studies focus on examining the impact of light recipes on plant growth, they solely concentrate on the light recipes used in their respective experiments and are unable to comment on the vast space of possible light combinations. Rather than solely relying on the comparison of preselected light recipes to identify the most effective one, our approach takes a different route. We leverage the available data to train a biomass estimator, which serves as a valuable tool for exploring the extensive landscape of light treatments and ultimately discovering the optimal light recipe.

The utilization of generative models for inverse design is a pertinent research direction, with several proposed methods focusing on designing devices in the domains of nanophotonics [11]–[14], polymer design [15] and design of complex metallic glasses [16]. PLNet stands out as the first study to employ generative techniques for the optimization of light recipes, specifically aiming to maximize biomass yield in plants. The utilization of deep neural networks for generating optimized light spectra based on desired growth outcomes represents a novel and promising approach in this field. This study aims to contribute to the growing body of research by exploring the capabilities of deep learning in the design and optimization of light recipes, ultimately improving crop yield, quality, and resource efficiency in indoor farming.

IV. PLANT LIGHT NETWORK: PLNET

In this work, we present PLNet (Plant Light Network), a neural network architecture designed to generate optimal light recipes for plant growth. Our approach involves training a generative neural network that has the capability to produce light recipes with a high probability of yielding high biomass. The architecture of PLNet is illustrated in Figure 2. The generator $G_\theta$ is trained to generate optimal light recipes, while
Fig. 2. The architecture of PLNet. Generator maps a random noise vector sampled from uniform distribution to the light recipe. Biomass estimator estimates biomass yield of the generated recipe. The generator is trained to generate optimal light recipe with a very high probability.

the biomass estimator $F_{\theta}$ is trained to estimate biomass of plants at a later stage based on the light recipe applied. First, the biomass estimator is trained using the plant growth data. Once the training of $F_{\theta}$ is completed, its weights are frozen, and it is used to estimate the biomass yield for different light recipes generated by $G_{\phi}$. By optimizing the generator we aim to generate light recipes that maximize the estimated biomass yield.

We represent the light recipe as a vector $x \in \mathbb{R}^N$, where $x = [x_1, x_2, \ldots, x_N]$ signifies a sequence of $N$ consecutive wavelengths. Each $x_i$ within this sequence corresponds to the PPFD of the $i^{th}$ wavelength. The search space of the light recipe, denoted as $S = \mathbb{R}^N$, defines the permissible range of PPFD values for each wavelength in the light recipe. To generate optimal light recipes, we introduce a random noise vector $z$ sampled from the uniform distribution $U([-1, 1])$ as an input to the generator. The generator, parameterized by $\phi$, maps $z$ to $x$, denoted as $G_{\phi}(z) = x$. In other words, the generator transforms noise vectors sampled from the uniform distribution into light recipe samples. We use $P_{\phi}$ to denote the light recipe distribution generated by $G_{\phi}$.

The objective of our approach is to maximize the probability of generating optimal light recipes that lead to high biomass yields. The biomass yield at the time of harvest for a specific light recipe $x$ is denoted as $b_x$. Therefore, our objective is to find the optimal parameter $\phi^*$ that maximizes the following integral:

$$\phi^* := \arg \max_{\phi} \int_S b_x \cdot P_{\phi}(x) \, dx$$  \hspace{1cm} (1)

However, evaluating Equation 1 over the entire search space $S$ is computationally infeasible. To overcome this challenge, we resort to approximating the objective by sampling a batch of light recipes $\{x^m\}_{m=1}^M$ from the recipe distribution $P_{\phi}$.

As our focus is on maximizing biomass yield at the time of harvest, we keep the age of the plant fixed at the harvest age $a^h$ when estimating biomass for the generated light recipe $x^m$. This leads to the following approximation:

$$\phi^* \approx \arg \max_{\phi} \frac{1}{M} \sum_{m=1}^M b_{x^m}$$  \hspace{1cm} (2)

In this equation, $M$ denotes the number of sampled light recipes, and the objective is to find the parameter configuration $\phi^*$ that maximizes the average biomass yield over the sampled recipes. This approximation enables a more computationally feasible approach to optimizing the generator for generating effective light recipes.

Furthermore, it is not practical to experimentally measure the biomass yield $b_x$ for each recipe in the selected sample batch in Equation 2. This is the reason we employ our biomass estimator, a neural network $F_{\theta}$ parameterized by $\theta$, to predict the biomass yields. The estimator receives the light recipe $x$ and the future age of the plant $a$ as input, predicting the biomass yield $b_{x,a}$ denoted as $F_{\theta}(x, a) = b_{x,a}$. The estimator $F_{\theta}$ is trained on a dataset that includes plant growth data at various ages until the harvest age, collected under diverse light recipes. This leads us to modify the objective function as follows:

$$\phi^* \approx \arg \max_{\phi} \frac{1}{M} \sum_{m=1}^M F_{\theta}(x^m, a^h)$$  \hspace{1cm} (3)

To ensure efficient optimization, we define a loss function $\mathcal{L}$ that, when minimized, maximizes the objective function.

$$\mathcal{L} = -\frac{1}{M} \sum_{m=1}^M F_{\theta}(x^m, a^h)$$
The gradient of the loss function is computed using the chain rule. It is important to note that $\theta$ is kept constant during the training of $G_\phi$. Furthermore, the age of harvest, $a_h$, is also fixed for a given plant.

$$\nabla_\phi L = \nabla_\phi x^m \cdot \nabla_x \phi f_0(x^m, a^h)$$

(4)

The gradient of the loss with respect to $\phi$ is calculated by multiplying the gradient of the generated light recipe $\nabla_\phi x^m$ with the gradient of the biomass estimator $\nabla_x \phi f_0(x^m, a^h)$.

This gradient calculation enables the optimization of the generator network $G_\phi$ by adjusting its parameters $\phi$ to generate light recipes that result in higher estimated biomass yields according to the fixed estimator $F_\theta$. By iteratively updating the parameters using gradient-based optimization techniques, the generator network learns to generate light recipes that lead to improved plant growth outcomes.

To facilitate the training process and ensure the generated light recipes are feasible, we integrate two distinct regularization terms aimed at ensuring the smoothness of the generated light recipe curve and controlling the total PPFD of the generated light recipe.

**Smoothness:** Light recipes often exhibit smooth curves. To encourage smoother light recipe curves, we introduce a smoothness regularization term. To compute the regularization value, we first compute $sx^m = [sx_1^m, sx_2^m, \cdots, sx_N^m]$, representing the smoother version of the generated recipe $x^m$. We define $k$ as the kernel size of the sliding window and compute smoothed values as follows

$$sx^m_i = \begin{cases} \frac{1}{k} \sum_{j=i}^{i+k-1} x^m_j, & \text{if } i \geq k \text{ and } i \leq N - k + 1, \\ x^m_i, & \text{otherwise}. \end{cases}$$

Finally, the smoothness loss is computed as follows

$$L_S = |x^m - sx^m|$$

**Total PPFD:** To control and limit the range of PPFD values within the generated light recipes, we incorporate a loss term when the sum of PPFD values exceeds a predefined threshold, denoted as $P_{\text{max}}$. This threshold represents the maximum allowable cumulative PPFD value for a given light recipe. The total PPFD loss is defined as follows

$$L_P = \begin{cases} 0, & \text{if } \sum_{i=1}^{n} x_i \leq P_{\text{max}}, \\ \sum_{i=1}^{n} x_i - P_{\text{max}}, & \text{otherwise}. \end{cases}$$

Here, $\sum_{i=1}^{n} x_i$ represents the total sum of PPFD values in the light recipe, and the loss is incurred only when this sum exceeds the specified threshold. This mechanism ensures that the generated light recipes adhere to a predetermined PPFD range, contributing to the feasibility and practicality of the generated solutions.

The overall loss function for training the generator becomes:

$$L_\phi = L + \lambda_s \cdot L_S + \lambda_p \cdot L_P$$

where $\lambda_s$ and $\lambda_p$ are weighting coefficients for the smoothness and total PPFD losses, respectively. By minimizing this loss function, we aim to find the optimal parameter $\phi^*$ that maximizes the biomass yield estimation.

### A. Biomass Estimator

Biomass estimator $F_\theta(x, a)$ plays a pivotal role in our methodology as it is tasked with predicting the biomass yield at the time of harvest, given a specific light recipe $x$ used for indoor plant growth. It is essential to note that the PPFD values of neighboring wavelengths in a light recipe are not entirely independent, and capturing the spatial relationships among these values is crucial for accurate predictions. To address this, we leverage a 1D convolutional neural network (1D CNN) to analyze and extract spatial features from the light recipe vector $[x_1, x_2, \cdots, x_N]$.

The primary objective of utilizing the 1D CNN is to extract meaningful features from the light recipe and map them to the corresponding biomass yield of the plant at the time of harvest. By leveraging the convolutional layers, the network can capture spatial dependencies and patterns in the light recipe, enabling a more comprehensive understanding of how different wavelengths and their interactions influence plant growth and biomass accumulation.

To train the biomass estimator $F_\theta$, we employ a dataset comprising plant growth data collected under various light recipes. These light recipes encompass a range of different spectral compositions and intensity levels. By associating the known biomass yields and age of the plant with their corresponding light recipes, we establish a supervised learning framework to train $F_\theta$. The training process allows the neural network to learn the complex relationships between the input light recipe vectors, the age of the plant and the resulting biomass yields. This learning facilitates accurate biomass estimation even for novel light recipes.

### B. Network Architecture

The architecture of the generative neural network in our PLNet framework is inspired by the Deep Convolutional Generative Adversarial Network (DCGAN) [17]. It consists of two fully connected layers, four transposed convolution layers, and a Gaussian filter at the end to remove small features. LeakyReLU activation functions are applied to all layers except for the final layer, which uses a tanh activation function. To enhance the diversity of the generated patterns, we incorporate dropout layers and batch normalization layers. These architectural choices allow the generative neural network to effectively capture the complex relationships between the input light recipe vectors and the desired light recipes, resulting in diverse and high-quality generated light recipes.

The biomass estimator network consists of four 1D CNN layers followed by a fully connected layer that connects to the output layer responsible for estimating the biomass. In alignment with the generative network, we incorporate dropout and batch normalization layers to enhance the model’s performance and generalization. The intermediate layers utilize the ReLU activation function, which introduces non-linearity to the network and enables better representation learning. On
In this study, we introduced PLNet (Plant Light Network), a novel approach for generating optimal light recipes to promote plant growth. PLNet utilizes a generative neural network, trained on plant growth data, to generate light recipes that are expected to result in high biomass yields. By formulating the

V. EXPERIMENTS

To evaluate the performance of PLNet in generating optimal light recipes for plant growth, we conducted experiments using plant growth data of Choy Sum (Brassica rapa var. parachinensis). Choy Sum is a leafy vegetable commonly cultivated in indoor farming systems.

A total of 41 distinct light treatments were employed in the generation of the training data, encompassing a range of total PPFD values spanning from 100 to 300 \( \mu \text{mol/m}^2\text{s} \). Out of these 41 light treatments, 25 were selected for training the models, while 8 were reserved for testing. The remaining 8 light treatments were exclusively used for validation purposes.

Concerning the training of the biomass estimator, multiple models were trained using 100 different seeds, and the model with the lowest validation error was identified as the optimal choice. A learning rate of 5e-3 was employed after exploring a range between 1e-2 and 1e-4. The batch size for training was set to 256, and the models were trained for 300 epochs. The training data was normalized to predict biomass in the range of 0-1.

The generator was trained for 1e4 epochs using a learning rate of 1e-4. During the training process, we conducted a hyperparameter search to determine the optimal values for \( \lambda_s \), \( \lambda_w \), \( k \), and \( P_{\text{max}} \). We explored a range of values for each hyperparameter: [0-100] for \( \lambda_s \), [0-1] for \( \lambda_w \), [3, 5, 7] for \( k \). For \( P_{\text{max}} \), we searched for [300, 400, 500, 600]. The search results led to the selection of \( \lambda_s = 20 \), \( \lambda_w = 0.001 \), \( k = 3 \), and \( P_{\text{max}} = 400 \) as the optimal values for these hyperparameters.

In our study, we performed a qualitative analysis of the light recipe generated by the generator. Specifically designed for Choy Sum, a leafy vegetable, the optimal light recipe was generated to maximize the growth and yield of Choy Sum plants. Figure 3 visually presents this optimized light recipe.

Detailed information on the light recipe is provided in Table I, including the Total PPFD and the corresponding PPFD values for specific color bands within the wavelength range. The estimated biomass for this recipe was found to be 0.99. It is crucial to emphasize that the biomass values have been normalized, where a value of 1.0 signifies the maximum attainable biomass yield. Additionally, it is noteworthy that, despite the PPFD range in the training data being 100-300 \( \mu \text{mol/m}^2\text{s} \), the generated light recipe demonstrates a PPFD of approximately 400 \( \mu \text{mol/m}^2\text{s} \) is optimal for Choy Sum.

Additionally, previous studies \([8], [19]–[21]\) have shown that specific ranges of light wavelengths have beneficial effects on plant growth. For instance, irradiation with UV-A light has been found to increase leaf area and biomass of plants \([8], [9]\), indicating the importance of including PPFD values in the UV-A range (315-400 nm). In our generated light recipe, we observe a PPFD of 23.96 \( \mu \text{mol/m}^2\text{s} \) in the range of 380-399 nm, suggesting that the generator has learned to incorporate UV-A light to enhance the biomass yield of Choy Sum.

Furthermore, far-red radiation (701-800 nm) has been shown to promote leaf area and biomass of plants \([19]–[21]\). Our generated light recipe includes a PPFD of 29.21 \( \mu \text{mol/m}^2\text{s} \) within the far-red range, indicating that the generator has also learned to incorporate far-red radiation to optimize biomass yield.

Overall, the qualitative analysis of the generated light recipe highlights the effectiveness of our approach in designing light recipes that promote plant growth. The generator has successfully learned to optimize the light spectrum by incorporating specific wavelengths known to enhance biomass yield in Choy Sum.

Further studies and experiments are needed to validate the performance of PLNet across different plant species, growth conditions, and indoor farming systems. Additionally, exploring the generalizability of PLNet to other plant growth parameters, such as nutrient uptake and leaf morphology, could provide valuable insights for holistic plant growth optimization.

VI. CONCLUSION

In this study, we introduced PLNet (Plant Light Network), a novel approach for generating optimal light recipes to promote plant growth. PLNet utilizes a generative neural network, trained on plant growth data, to generate light recipes that are expected to result in high biomass yields. By formulating the
light recipe as a vector and using a generative neural network, we aim to maximize the probability of generating optimal light recipes.

To address the challenge of evaluating the objective function over the entire search space, we employed a sampling-based approximation approach. By sampling a batch of light recipes from the generator, we obtained an estimate of the objective function, allowing us to optimize the generator network accordingly.

Furthermore, to estimate the biomass yield associated with each generated light recipe, we developed a biomass estimator network. This network consists of multiple 1D CNN layers followed by a fully connected layer and a sigmoid activation function, allowing us to optimize the generator network accordingly.

In summary, our work makes a valuable contribution to the field of indoor plant cultivation by introducing a neural network-based framework, PLNet, for generating optimal light recipes. Such experimental studies will provide a more rigorous evaluation of the performance and effectiveness of the generated light recipes.

ACKNOWLEDGEMENT

This work was supported by the Agency for Science, Technology and Research (A*STAR) under its Industry Alignment Fund - Pre Positioning (IAF-PP) (A19D9a0096).

REFERENCES


