IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 10 June 2024, accepted 24 June 2024, date of publication 4 July 2024, date of current version 23 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3423455

RESEARCH ARTICLE

Lettuce Fresh Weight Prediction in a Plant Factory Using Plant Growth Models

YUYA HOSODA[1], (Member, IEEE), TOTA TADA[2], AND HITOSHI GOTO[2,3]
[1]Graduate School of Engineering Science, Osaka University, Osaka 560-8531, Japan
[2]Graduate School of Computer Science and Engineering, Toyohashi University of Technology, Aichi 441-8580, Japan
[3]Information and Media Center, Toyohashi University of Technology, Aichi 441-8580, Japan

Corresponding author: Yuya Hosoda (hosoda.yuya.ho@tut.jp)

This work was supported by the Naito Research Grant.

ABSTRACT This paper proposes a novel method to predict the fresh weight of lettuce at the shipping stage in a plant factory using the early-stage growth images. It is well-established that the size and shape of plants correlate with their fresh weight. The proposed method captures chlorophyll fluorescence-based growth images daily and extracts geometric features such as projection area, edge length, and skeleton length. We design a regression model to predict the fresh weight using the dimensionality-reduced historical features. However, without considering growth statuses, the dimensionality reduction approach leads to decreased predictive performance for mature and slower-growing plants. In this paper, we generate a plant growth model that simulates the growth process by integrating multiple growth records based on the comparison of growth statuses. The proposed method then reduces the dimensionality by fitting historical features to the plant growth model to obtain future features. Experimental results demonstrate that the proposed method accurately predicts the fresh weight and achieves the coefficient of determination of 0.885, root mean square error of 8.790 g, and mean absolute error of 6.684 g when predicting the fresh weight ten days ahead using growth images from the past ten days.

INDEX TERMS Plant factory, growth prediction, machine learning, curve fitting, regression model.

I. INTRODUCTION

The term ''plant factory'' refers to the agricultural technology for cultivating plants within enclosed container facilities using advanced information and communication technology [1], [2]. These facilities are equipped with artificial lighting and air conditioning systems, enabling stable year-round harvests irrespective of external challenges such as unfavorable weather or pests. A notable feature is hydroponic cultivation, which involves circulating liquid fertilizer in the facility, facilitating the growth of leafy vegetables even in arid areas. However, the continuous operation of these systems leads to higher cultivation costs of plant factories compared to traditional open-field cultivation.

This study addresses a critical challenge in plant factory operations: optimizing cultivation environments to reduce operational costs. Room temperature [3], [4] and liquid fertilizer concentration [5], [6] are crucial factors that determine nutrient-rich plant production. Airflow management is also vital for preventing leaf scorch in leafy vegetables [7], [8]. Since electricity accounts for a significant portion of cultivation costs, efficient system operation is paramount [9], [10]. Consequently, developing an operational plan that minimizes energy consumption has become the focus in this field.

Artificial lighting, a vital component of the cultivation control system, is essential for plant growth, but it consumes much electrical energy. Selecting the appropriate light wavelength [11], [12] and intensity [13], [14] can enhance the nutritional value of plants. Adjusting light exposure duration influences the circadian rhythms of a plant [15], [16]. By leveraging the market price forecasting technique [17], [18], growers can align growth cycles with market demands, thereby reducing food waste due to overproduction. The operating plan for artificial lighting must be carefully crafted to balance plant growth and energy efficiency.

The associate editor coordinating the review of this manuscript and approving it for publication was Joewono Widjaja.

There are various methods for analyzing the plant growth. Photogrammetry [19], [20], which generates 3D models from images taken from multiple angles, allows for non-contact size measurement. Using historical environmental data, researchers estimate electrical conductivity and evapotranspiration in soilless culture systems, which indicate the physiological activity of plants [21], [22]. Image recognition techniques automatically detect diseases and pests from growth images [23], [24]. While these methods aim to measure the current plant growth, the future growth prediction will enable to optimize long-term cultivation in plant factories.

Plant growth prediction methods analyze historical growth records using machine learning to obtain the future plant growth, such as leaf shape [25], [26] and plant height [27], [28]. Fresh weight for vegetables is crucial as a criterion for determining market prices. Liu et al. [29] predict the fresh weight of lettuce based on phenotypic and environmental data. Minchin et al. [30] obtain the future fresh weight of fruits by fitting size measurements to mathematical functions. Meanwhile, growth images are utilized to analyze plant structures related to the fresh weight [31], [32], [33]. Kim et al. [34] show that geometric features in RGB color-based growth images correlate with the plant growth. Nagano et al. [35] reduce the dimensionality of the historical features using principal component analysis (PCA) to predict the future fresh weight of lettuce.

Our previous research [36] utilizes chlorophyll fluorescence-based growth images to predict the fresh weight of lettuce. These images capture only the light energy emitted by plants [37], [38], [39], so the edge length and projection area can be measured without preprocessing. Since plants grow according to a mathematical model [34], [40], this method converts historical features to future features by fitting them into a mathematical function to reduce the dimensionality. A multiple linear regression (MLR) model with limited growth records achieves the fresh weight prediction. However, since the mathematical functions are directly designed from the data points, measurement errors and insufficient data degrade prediction performance. Additionally, the feature conversion does not consider the different growth statuses of plants, such as mature and slower-growing plants [41].

This paper proposes a novel approach using plant growth models to predict the fresh weight of lettuce. Three notable advancements distinguish our previous research [36]. First, we incorporate skeleton length as an additional geometric feature, recognizing the horizontal spread of leaves. Second, the proposed method fits historical features to a plant growth model that simulates the plant growth process and obtains future features while reflecting information about the growth status. We also develop a reliable plant growth model by merging multiple growth records based on comparing growth statuses, thereby suppressing the effects of measurement errors and insufficient data. Third, we validate the predictive performance through leave-one-out cross-validation

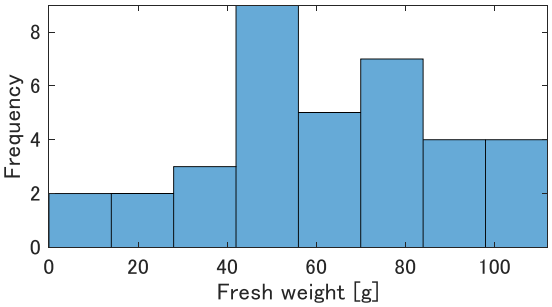

FIGURE 1. Histogram for the fresh weight of lettuce.

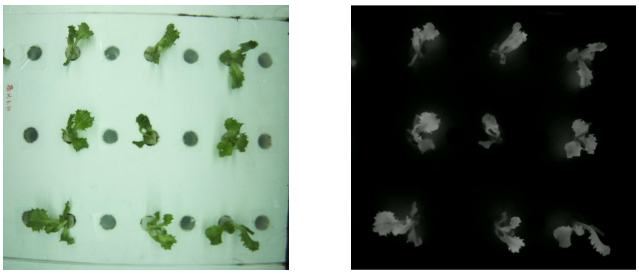

FIGURE 2. RGB color-based growth image (Left) and chlorophyll fluorescence-based growth image (Right).

on the growth records of 36 plants at various growth statuses.

The contributions of this paper are summarized as follows:

- We evaluate the predictive performance of the proposed method compared to traditional methods [35], [36]. The proposed method outperforms them with a coefficient of determination (R2) of 0.885, root mean square error (RMSE) of 8.790 g, and mean absolute error (MAE) of 6.684 g for the fresh weight prediction ten days ahead using growth images from the past ten days.
- We discuss parameter settings for the proposed method. All geometric features with the feature conversion using the plant growth model contribute to accurate fresh weight prediction independent of the growth status.
- We examine mathematical functions to verify the feature conversion. The feature conversion using higher-dimensional mathematical functions, which include inflection points, improves the predictive performance, suggesting that the plant growth rate varies depending on the growth status.

This paper is structured as follows. In Section II, we introduce the cultivation environment and outlines the proposed fresh weight predictive method using plant growth models. The experimental results are discussed in Section III. Finally, Section IV presents our conclusions.

II. MATERIALS AND METHODS

A. CULTIVATION ENVIRONMENT

This paper aims to predict the future fresh weight of nontuberous lettuce in a plant factory. We chose the Frillice lettuce, originating from Chile. The facilities regulated the cultivation conditions within an enclosed plant factory

at the Toyohashi University of Technology, Japan. Water temperatures were maintained in the range of 25°C–28°C, and the carbon dioxide concentration was kept within a range of 1000–1100 ppm. An air-conditioning system consistently maintained the ambient temperature at 25°C.

The growth rack had an artificial lighting setup, combining blue/red and white LED lights. The lighting regime comprised 16 h of light and 8 h of darkness. We calibrated the light intensity to 58 W/m^2 for the red/blue LEDs and 76 W/m^2 for the white LEDs to optimize the conditions for lettuce growth. Additionally, the nutrient solution was circulated throughout the facility, moving through filters, tanks, and pumps. The electrical conductivity of the liquid fertilizer was controlled between 1.00 and 1.30 mS/m, with the pH levels maintained steady between 6.92 and 6.95.

In this study, we cultivated 36 plants from June 30, 2021, to August 2, 2021. Seeds were sown on sponges, which were shielded from light for the first two days. After 14 days, the seedlings were moved to the growing rack at intervals of 15 cm, and this day is regarded the first day of planting ($t = 1$). We measured the fresh weight of lettuce without roots on $t = 20$. Figure 1 shows the histogram for the fresh weight of 36 plants. The dataset included plants at various growth statuses, where the maximum, minimum, median, and average fresh weights were 106.52, 8.86, 61.57, and 62.58 g, respectively.

This paper focuses on the photosynthetic process, which transforms light energy into chemical energy to produce carbohydrates. Chlorophyll within chloroplasts synthesizes glucose from sunlight, a crucial energy source for plant growth and survival. Plants emit unused light energy at night, and this phenomenon is known as chlorophyll fluorescence. This paper used the Raspberry Pi Camera Module V2 to capture chlorophyll fluorescence. We recorded a video with a capture duration of 30 seconds, a frame size of 1640 × 1232 pixels, and a frame rate of 20 frames per second at a stabilization period of 8 hours of darkness. We mounted a long-pass filter (SC-66; FUJIFILM) on the camera lens to filter out wavelengths shorter than 660 nm, including visible and ultraviolet light. This filter allows the transmission of infrared light with wavelengths greater than 680 nm, which is emitted by plants. The frame captured at 10 seconds was selected as the chlorophyll fluorescence-based growth image. Additionally, we employed a wide-angle lens (RP-L165; Entaniya) with a field of view of 73 degrees (vertical), 134 degrees (horizontal), and 160 degrees (diagonal) to capture multiple plants simultaneously. We corrected lens distortion by a calibration technique involving a chessboard pattern.

Figure 2 presents the RGB color-based and chlorophyll fluorescence-based growth images. RGB color-based growth images, which were captured using the Raspberry Pi Camera Module V2 without the long-pass filter, contain obstacles, necessitating image preprocessing to remove them. Such preprocessing incurs additional costs and labor. In contrast, chlorophyll fluorescence-based growth images, which only contain plants, allow the direct extraction of features without preprocessing. This work utilizes chlorophyll fluorescence-based growth images, rather than RGB color-based growth images, for fresh weight prediction. Since plants emit chlorophyll fluorescence as a red light at night, we took only one growth image per day from $t = 1$ to $t = 20$. We then divided it into 320 × 240 pixels so that each plant was centered. However, after $t = 13$, the leaves of all plants overlap and cannot be divided accurately. In this paper, we utilize $t = 1, \ldots, T$-th ($T < 13$) growth images.

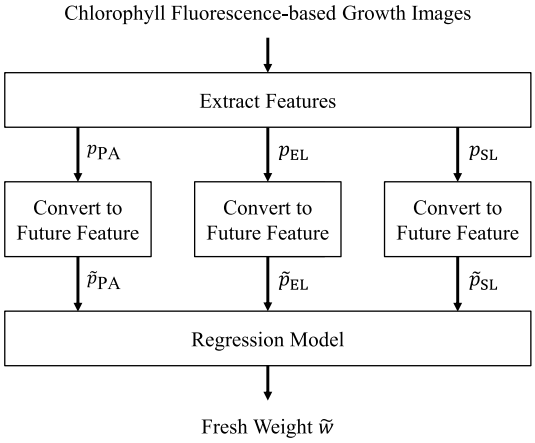

FIGURE 3. Block diagram of the fresh weight prediction method using plant growth models.

B. FRESH WEIGHT PREDICTION METHOD USING PLANT GROWTH MODELS

Figure 3 shows a block diagram of our proposed method for predicting the flesh weight of lettuce with plant growth models. Initially, we extract the key features, namely, projection area, edge length, and skeleton length, from chlorophyll fluorescence-based growth images. The historical features are then integrated into plant growth models to obtain future features. Finally, a regression model predicts the fresh weight of lettuce at the shipping stage.

1) FEATURE EXTRACTION FROM GROWTH IMAGES

We used the projection area to measure plant size. In chlorophyll fluorescence-based growth images, areas of active photosynthesis are more luminous and represent plant areas. We extracted the plant area by Otsu's method. Figure 4a shows the extracted plant area. Because of the overshadowing of internal leaves by outer ones, some areas are missed, resulting in a projection area smaller than the actual size. Considering the expansion pattern of lettuce leaves, which typically grow from the inside out, areas currently obscured may have been visible in previous growth images. Therefore, our method combines the current and previous growth images to interpolate the missing areas. Let $B_t \in (0, 1)$ denote a binary image. The interpolated binary image, $\hat{B}_t$, is defined as follows:

$$\hat{B}_t = B_t \cup B_{t-1}. \tag{1}$$

(a) (b)

FIGURE 4. Plant area in chlorophyll fluorescence-based growth images. (a) Without interpolation. (b) With interpolation.

Figure 4b shows the interpolated plant area. The missing areas are filled so that the projection area can be accurately represented. We then define the projection area as the ratio of the plant area to total area of the growth image.

Frillice lettuce shows a growth pattern wherein its leaves develop from the inside outward, resulting in an increased overlap of leaves as the plant matures. That is, the structural complexity escalates with plant growth. To analyze the complexity, we extract leaf contours by the Canny edge detection method, which employs two threshold values to identify reliable edges. Additionally, the proposed method measures the plant structure by calculating the skeleton length. The skeleton length is a continuous line that outlines the plant area minimally, capturing the structural changes that occur during plant growth.

2) FEATURE CONVERSION USING PLANT GROWTH MODELS

The proposed method arranges the geometric features from growth images in historical order to predict future fresh weight. Because of high-dimensional features, the predictive model should reduce their dimensionalities to mitigate model complexity. It is well known that leafy plants like lettuce follow an exponential growth model [34], [40]. The proposed method obtains future features while considering the growth status by fitting the historical features to a plant growth model that simulates the plant growth process. We approximate the multiple growth records with mathematical functions to generate the plant growth model, suppressing the effects of measurement errors and insufficient data. However, since each plant has a different growth status, it cannot directly merge growth records. The proposed method compares their growth statuses using mean squared error (MSE) and merges the growth records in order while shifting the day variable. Note that there is a high chance of error propagation over time due to the lack of ground truth. Therefore, the proposed method generates multiple plant growth models by switching the combination order randomly and selects the model with the best fit with the growth records.

Algorithm 1 outlines the process for generating a plant growth model. The proposed method generates N plant growth models using historical features of M growth records. The n-th ($n = 1, \ldots, N$) plant growth model, g_n, is formed by fitting historical features to a mathematical function, f. Let $\mathbf{P}[m]$ ($m = 1, \ldots, M$) represent a multidimensional data for the m-th plant, which includes historical features from growth images over T days. We define a set of feature variables $\mathbf{y}$ and a set of day variables $\Delta\mathbf{t}$. For the first plant, day variables are assigned using Permute(1, T), creating a sequence from 1 to T. The function, Fit(f, $\Delta\mathbf{t}$, $\mathbf{y}$), calculates the function parameters, minimizing the distance between function f and data point ($\Delta\mathbf{t}$, $\mathbf{y}$). In subsequent plants, the proposed approach iteratively generates the plant growth model and places data points by reading growth records sequentially.

Data points for the subsequent plant are placed by comparing the growth status, such as mature or slower-growing, with the first plant based on the plant growth model. By adjusting the day variables, our method seeks the best fit between the plant growth model and historical features. The shift, s, ranges from s_{start} to s_{end}, increasing by Δs. In this study, we set $s_{\text{start}} = -5$, $s_{\text{end}} = 5$, $\Delta s = 0.1$. Negative

Algorithm 1 Plant Growth Model Generation Algorithm

Require: $\mathbf{P}, T, f, s_{\text{start}}, s_{\text{end}}$
Ensure: g

$\text{MSE}_{\text{min}} \leftarrow \infty$
$g \leftarrow$ null
for $n \leftarrow 1$ to N **do**
  $\mathbf{P}^* \leftarrow$ Randomize($\mathbf{P}$)
  $\Delta\mathbf{t} \leftarrow$ Permute(1, T)
  $\mathbf{y} \leftarrow \mathbf{P}^*[1]$
  $g_n \leftarrow$ Fit(f, $\Delta\mathbf{t}$, $\mathbf{y}$)
  for $m \leftarrow 2$ to M **do**
    $s_{\text{opt}}, s \leftarrow s_{\text{start}}$
    $\Delta\mathbf{t}_n \leftarrow$ Permute($s+1, s+T$)
    $\mathbf{y}_n \leftarrow \mathbf{P}^*[m]$
    $\text{MSE}_{\text{opt}} \leftarrow$ CalcMSE(g_n, $\Delta\mathbf{t}_n$, $\mathbf{y}_n$)
    while $s \leq s_{\text{end}}$ **do**
      $s \leftarrow s + \Delta s$
      $\Delta\mathbf{t}_n \leftarrow$ Permute($s+1, s+T$)
      $\text{MSE}_n \leftarrow$ CalcMSE(g_n, $\Delta\mathbf{t}_n$, $\mathbf{y}_n$)
      if $\text{MSE}_n < \text{MSE}_{\text{opt}}$ **then**
        $\text{MSE}_{\text{opt}} \leftarrow \text{MSE}_n$
        $s_{\text{opt}} \leftarrow s$
      end if
    end while
    $\Delta\mathbf{t} \leftarrow$ Append($\Delta\mathbf{t}$, Permute($s_{\text{opt}}+1, s_{\text{opt}}+T$))
    $\mathbf{y} \leftarrow$ Append($\mathbf{y}$, $\mathbf{P}^*[m]$)
    $g_n \leftarrow$ Fit(f, $\Delta\mathbf{t}$, $\mathbf{y}$)
  end for
  $\text{MSE}_n \leftarrow$ CalcMSE(g_n, $\Delta\mathbf{t}$, $\mathbf{y}$)
  if $\text{MSE}_n < \text{MSE}_{\text{min}}$ **then**
    $g \leftarrow g_n$
    $\text{MSE}_{\text{min}} \leftarrow \text{MSE}_n$
  end if
end for

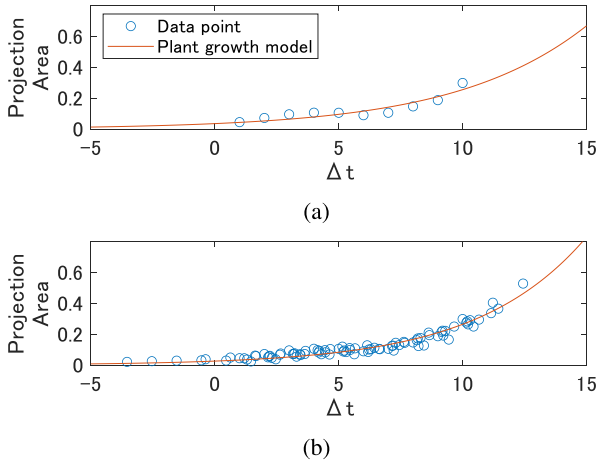

FIGURE 5. Plant growth models for the projection area. We set the parameter with $M = 9$, $T = 10$ and the Logistic function. (a) $g_1(\Delta \mathbf{t})$. (b) $g_{10}(\Delta \mathbf{t})$.

values indicate that the current plant growth lags relative to previous plants and vice versa. In this paper, we evaluate the fitting quality using MSE:

$$\text{CalcMSE}(g, \Delta \mathbf{t}, \mathbf{y}) = ||g(\Delta \mathbf{t}) - \mathbf{y}||^2. \tag{2}$$

The algorithm identifies the shift s_{opt} that minimizes the MSE as the optimal shift. Shifted day variables and features are incorporated into the original sets through Append, which adds new data points to the dataset, updating the plant growth model. The update process continues up to the M-th plant. The algorithm produces N plant growth models by randomly changing plant indices with Randomize($\mathbf{P}$). The model with the lowest MSE with the total features is then chosen as the final plant growth model g.

Figure 5 illustrates the plant growth model for the projection area. We set the parameter with $M = 9, T = 10$ and the Logistic function. When multiple growth records are placed with shifting day variables relative to growth statuses, the data points follow a typical growth process. These results are consistent with previous research [34], [40] that showed that plant growth could be represented using a mathmatical function. Therefore, by approximating the historical features with mathematical functions, we can obtain a plant growth model that simulates plant growth with limited growth records. Additionally, the plant growth model varies with the number of growth records. Approximating a single growth record by fixing the day variable is sensitive to measurement error and insufficient data. This paper thus uses multiple growth records to generate plant growth models for geometric features.

The proposed method obtains future features from the historical features using the plant growth model. Let $\tilde{T}$ denote the date at the shipping stage. Algorithm 2 outlines the procedure for calculating future features. Let $\mathbf{p}$ denote the historical features of the plant to be predicted for the fresh weight. Initially, the proposed method adjusts the day variable by s_{opt} to align with the plant growth model, using the same process as that employed in plant growth model generation.

Algorithm 2 Future Feature Calculation Algorithm

Require: $\mathbf{p}$, T, $\tilde{T}$, g, s_{start}, s_{end}
Ensure: $\tilde{p}$

```
s ← s_start
Δt ← Permute(s + 1, s + T)
MSE_opt ← CalcMSE(g, Δt, p)
s_opt ← s
while s ≤ s_end do
    s ← s + Δs
    Δt ← Permute(s + 1, s + T)
    MSE ← CalcMSE(g, Δt, p)
    if MSE < MSE_opt then
        MSE_opt ← MSE
        s_opt ← s
    end if
end while
p̃ ← CalcFutureFeature(g, s_opt, T̃)
```

The future feature $\tilde{p}$ is then calculated using the function CalcFutureFeature(g, s_{opt}, $\tilde{T}$), as follows:

$$\tilde{p} = g(s_{\text{opt}} + \tilde{T}). \tag{3}$$

For early-stage plants, a positive s_{opt} leads to a larger $\tilde{p}$, whereas for slower-growing plants, a negative s_{opt} results in a smaller $\tilde{p}$. Thus, our method efficiently reduces the dimensionality of geometric features while accurately reflecting information about the growth status.

3) FRESH WEIGHT PREDICTION USING REGRESSION MODELS

The proposed method predicts the fresh weight of lettuce at the shipping stage by using future features in an MLR model. Let $\tilde{w}$ represent the predicted fresh weight. The MLR model is formulated as follows:

$$\tilde{w} = a_{\text{PA}} \cdot \tilde{p}_{\text{PA}} + a_{\text{EL}} \cdot \tilde{p}_{\text{EL}} + a_{\text{SL}} \cdot \tilde{p}_{\text{SL}} + b, \tag{4}$$

where $\tilde{p}_{\text{PA}}$, $\tilde{p}_{\text{EL}}$, and $\tilde{p}_{\text{SL}}$ denote the future feature for the projection area, edge length, and skeleton length, respectively. These future features are derived from the plant growth models described in Algorithm 2. The parameters a_{PA}, a_{EL}, a_{SL}, and b are the coefficients of the MLR model and are determined by the least squares method. This method aims to minimize the sum of the squares of the discrepancies between the observed fresh weights and those predicted by the model across all samples in the training data.

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

We conducted a series of experiments to assess the efficacy of our proposed method for predicting the fresh weight of lettuce. The independent variables of the regression model were the future features for the projection area (PA), edge length (EL), and skeleton length (SL). We adopted four mathematical functions: quadratic, cubic, quartic, and

TABLE 1. Predictive performance of various combinations of mathematical functions, including quadratic (Quad), cubic, quartic, and logistic functions. From the 64 possible combinations, the top five and bottom five in terms of R2 are shown in the table in its upper and lower sections, respectively.

Mathematical Function			Metrics		
PA	EL	SL	R2	RMSE [g]	MAE [g]
Logistic	Logistic	Quartic	0.885	8.790	6.684
Quartic	Logistic	Logistic	0.877	9.084	7.049
Logistic	Logistic	Cubic	0.871	9.284	7.247
Cubic	Cubic	Quartic	0.869	9.355	7.452
Quartic	Logistic	Quartic	0.869	9.375	7.039
Logistic	Quartic	Quad	0.832	10.619	7.852
Quad	Logistic	Quad	0.830	10.665	7.931
Quad	Cubic	Quad	0.829	10.692	7.932
Cubic	Quartic	Quad	0.827	10.784	8.058
Quad	Quad	Quad	0.825	10.822	8.243

Logistic functions. The Logistic function is expressed as follows:

$$f(t) = \frac{a}{1 + \exp(-a \cdot c \cdot (t - b))}, \tag{5}$$

where a represents the maximum value of the function; b shifts the curve along t; c controls the steepness or slope of the curve. We set default parameters for the proposed method with $M = 9, N = 10, T = 10, \tilde{T} = 20$.

We assessed the predictive performance of the proposed method using a leave-one-out cross-validation method with $K = 36$ plants. In each cross-validation cycle, nine plants were used to generate the plant growth model, and 26 plants were used to design the weight prediction model. One plant was set aside for testing. We evaluated the predictive performance using several metrics: coefficient of determination (R2), root mean square error (RMSE), and mean absolute error (MAE). These metrics are defined as follows:

$$\mathrm{R2} = 1 - \frac{\sum_{k=1}^{K}(w_k - \tilde{w}_k)^2}{\sum_{k=1}^{K}(w_k - \bar{w})^2}, \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(w_k - \tilde{w}_k)^2}, \tag{7}$$

$$\mathrm{MAE} = \frac{1}{K}\sum_{k=1}^{K}|w_k - \tilde{w}_k|, \tag{8}$$

where w_k denotes the actual fresh weight for k-th plant, and $\bar{w}$ denotes the average of the actual fresh weight.

First, we examined the efficacy of various mathematical functions in generating plant growth models. Table 1 summarizes the predictive performance of different combinations of mathematical functions. From the 64 possible combinations, the table highlights the top five and bottom five in terms of R2 in its upper and lower sections, respectively. A comparison between the best and worst performers shows significant discrepancies: the differences in R2, RMSE, and MAE are 0.060, 2.032, and 1.559, respectively, implying the impact of the combination of mathematical functions on predictive performance. In the effective combinations, the cubic function

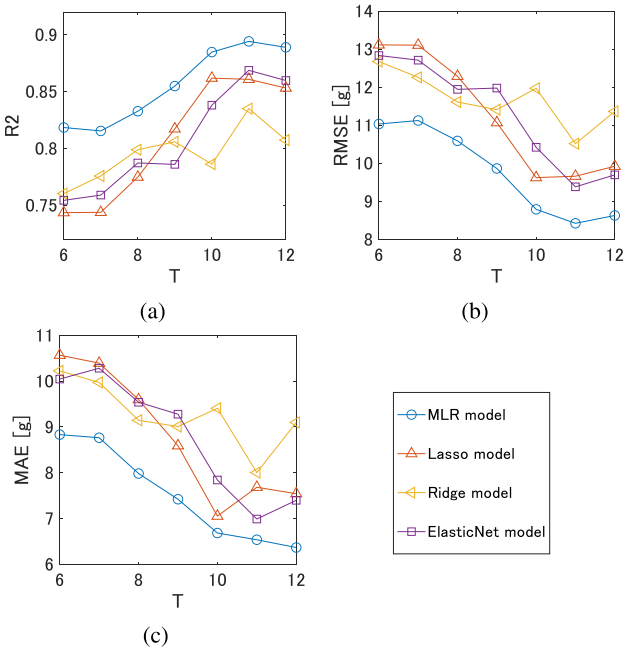

FIGURE 6. Predictive performance with the length of the historical features. The compared regression models include MLR (Ours), Lasso, Ridge, ElasticNet models. (a) R2. (b) RMSE. (c) MAE.

was present thrice, quartic functions appeared five times, and logistic functions appeared seven times. Higher-dimensional mathematical functions with inflection points are more adept at representing the plant growth process. Conversely, the less effective combinations consistently included the quadratic function. These results show that plant growth rates vary with the growth status, necessitating more sophisticated modeling approaches. In the following experiments, the logistic function is used for the feature conversion with projection area and edge length, while the quartic function with skeleton length.

We investigated the the predictive performance with respect to the length of the historical features and regression models. A shorter length of the historical features indicates an earlier prediction of fresh weight at the shipping stage. The proposed method, which uses an MLR model, was compared with other models incorporating feature selection regularization techniques such as Lasso, Ridge, and Elastic Net models. The regularization weight was optimized through 5-fold cross-validation, with the MSE as the loss function. Figure 6 shows the predictive performance of each model with the length of the historical features. Predictive accuracy improves across all metrics as the predictive date approaches the shipment date. These results suggest that extracting more geometric features from growth images enables reliable feature conversion using the plant growth model–however, the benefit of the length of the historical features stagnated around $T = 10$. The proposed method was challenging to accurately extract geometric features from growth images after $t = 11$ because of leaf overlapping between plants. Traditional approaches [34],

TABLE 2. Predictive performance of feature combinations, including projection area, edge length, and skeleton length.

Features			Metrics		
PA	EL	SL	R2	RMSE [g]	MAE [g]
✓	-	-	0.830	10.685	8.205
-	✓	-	0.811	11.265	8.664
-	-	✓	0.858	9.766	7.357
✓	✓	-	0.818	11.051	8.733
✓	-	✓	0.857	9.806	6.798
-	✓	✓	0.863	9.592	6.519
✓	✓	✓	0.885	8.790	6.684

[35] individually take growth images for each plant or limit the analysis area. Given that actual plant factories often reduce the spacing between plants to maximize area yield, current predictive methods have not developed effective measures against leaf overlapping between plants. Introducing image recognition techniques [42], [43] will enhance the performance of extracting the plant region from the leaf overlapping growth image. Additionally, the MLP model consistently outperformed the other models, regardless of the length of historical features. These results indicate that using a comprehensive set of available features, as opposed to selective feature extraction, is crucial for the accuracy of the predictive model.

Table 2 summarizes the predictive performance of the feature combinations. When employing a single feature, the predictive model using the skeleton length achieved the highest accuracy, with an R2 of 0.858, RMSE of 9.766, and MAE of 7.357. The skeleton length, indicative of plant structure, was a reliable predictor of weight because mature-growing plants exhibit horizontal leaf extension at the seedling stage. The projection area, while indicative of plant size and its correlation to fresh weight, was less effective because growth conditions near the base were not adequately captured in top-down growth images. The edge length, although helpful, could not accurately predict fresh weight because of minimal leaf overlap at the early stage. The feature combination with the edge length and skeleton length improved the prediction performance compared to the skeleton length alone, with an R2 of 0.005, RMSE of 0.174, and MAE of 0.838. These results highlight that edge length contributes valuable information that is not captured by skeleton length. Conversely, the combination of the projection area and the skeleton length slightly reduced R2 by 0.001 and increased RMSE by 0.040 but improved MAE by 1.407. Despite some outliers, this combination enhanced the predictive performance, confirming the importance of the projection area in fresh weight prediction. Finally, the predictive model that integrated all three features yielded the best results, with an R2 of 0.885 and RMSE of 8.790. These findings demonstrate all geometric features have their own important role in predicting fresh weight.

We discussed on the parameters for the plant growth model. Table 3 lists the predictive performance of the proposed method under different parameter settings for generating plant growth models. When the plant growth model was

TABLE 3. Predictive performance of methods using plant growth models under different parameters, including the number of the plants, the length of the historical features, and the number of plant growth models to be generated.

Model Parameters			Metrics		
M	T	N	R2	RMSE [g]	MAE [g]
1	10	10	0.733	13.382	10.977
5	10	10	0.870	9.337	7.623
9	10	10	0.885	8.790	6.684
9	6	10	0.812	11.225	8.950
9	8	10	0.857	9.776	7.052
9	10	10	0.885	8.790	6.684
9	12	10	0.867	9.446	7.696
9	10	1	0.874	9.199	7.052
9	10	4	0.880	8.981	6.976
9	10	7	0.878	9.032	7.126
9	10	10	0.885	8.790	6.684

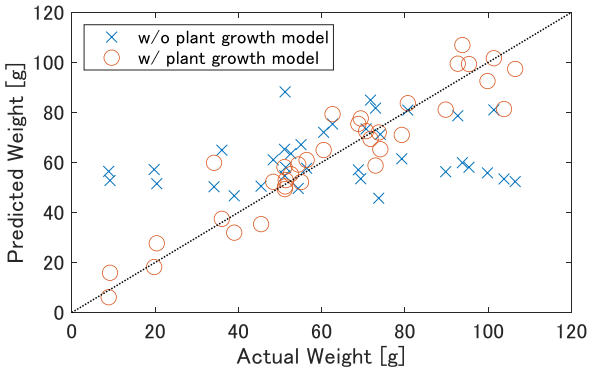

FIGURE 7. Results of fresh weight prediction. The "×" and "o" marks represent the predictive methods using the feature conversion without or with the plant growth model.

generated from multiple growth records with $M = 9$, it improved the R2 to 0.152, reduced the RMSE to 4.592, and decreased the MAE to 4.293 compared to using a single growth record with $M = 1$. Merging multiple growth records can mitigate the impact of measurement errors in generating a reliable plant growth model. When the number of growth records was fixed at $M = 9$, all metrics improved as the length of historical features increased from $T = 6$ to $T = 10$. A limited number of the historical features hinders the plant growth model from reflecting the growth status accurately. There was a decline in predictive performance at $T = 12$ because it can be attributed to leaf overlapping in growth images beyond $t = 11$, leading to measure errors. For generating a high-quality plant growth model, it is crucial to accurately extract the geometric features from growth images. When a single plant growth model was generated with $N = 1$, the MAE was above 7.000 g because it can be affected by the propagation of measurement errors. On the other hand, when the one that best fits the data points was selected among multiple plant growth models, the proposed method achieved the highest predictive performance at $N = 10$. Therefore, generating multiple iterations of the plant growth model contributes to the enhancement of predictive performance.

We evaluated the effectiveness of the feature conversion using plant growth models. Figure 7 shows the results of fresh weight prediction. The "×" and "o" marks represent

TABLE 4. Comparison of predictive performance using different feature conversion approaches and regression models. The first combination corresponds to Nagano et al.'s method [35], while the fourth combination corresponds to the proposed method.

Feature Conversion	Regression Model	Metrics R2	RMSE [g]	MAE [g]
PCA	GBR	0.841	10.333	7.868
PCA	MLP	0.817	11.073	8.511
PGM	GBR	0.743	13.138	10.012
PGM	MLP	0.885	8.790	6.684

the predictive methods using the feature conversion without or with the plant growth model. Our previous method [36], directly approximating data points, achieved satisfactory results for a median weight of 61.57 g, but it tended to overestimate the weight of slower-growing plants and underestimate that of mature-growing plants. In contrast, the proposed method, using feature conversion with plant growth models, delivered accurate weight predictions across various growth statuses. Therefore, the plant growth model is effective in predicting fresh weight of lettuce independent of growth status.

Finally, we compared the predictive performance of the proposed method with a baseline. As the baseline, we adopted the method devised by Nagano et al. [35], which uses PCA for the feature conversion and the gradient-boosting regression (GBR) model. We also prepared combinations of the feature conversion methods and regression models. In this paper, we set PCA up to the third principal component to match the dimensionality of features obtained by the plant growth model (PGM). Additionally, we optimized the GBR model using the least squares boosting algorithm with a learning rate of 0.1. Table 4 shows the comparison of predictive performance. In the methods utilizing PCA, which reduces the dimensionality without considering the growth status, the RMSE exceeded 10 g regardless of the regression models. When the historical features were converted using the plant growth model considering the growth statuses, the MLP model achieved the best values across all evaluation metrics. Since PCA transforms features to maximize variance, it was suitable for the GBR model, which could capture nonlinear relationships. The feature conversion using PGM represented the future feature larger for mature-growing plants and smaller for slower-growing plants. Hence, the MLP model, which designed the linear representation, was superior to the GBR model. Since actual plant factories cultivate plants with various growth statuses simultaneously, the proposed method using the plant growth and MLP models is practical.

IV. CONCLUSION

This paper introduced a novel method for predicting the fresh weight of lettuce at the shipping stage using plant growth models. The experimental results showed that the proposed method achieved high predictive performance, with an R2 of 0.885, RMSE of 8.790, and MAE of 6.684, when predicting the fresh weight ten days ahead using growth images from the past ten days. Through multiple comparative experiments, we demonstrated that plant growth models enhanced the predictive performance for plants at different growth statuses, such as mature and slower-growing plants. Since the growth of leafy vegetables follows a mathematical model, our approach is valid for plants other than lettuce. It should be noted, however, that measure errors due to leaf overlapping between plants degrade prediction performance. Additionally, our research revealed that the feature conversion using higher-dimensional mathematical functions outperformed using simple mathematical functions, so the plant growth changes at each growth status. Future work controls plant growth in actual plant factories using the proposed weight prediction method and assesses growth conditions based on the plant growth model.

ACKNOWLEDGMENT

Yuya Hosoda was with the Center for IT-Based Education, Toyohashi University of Technology, Aichi, Japan. The authors wish to acknowledge Kotaro Takayama at Toyohashi University of Technology, for providing data on the plant factory.

REFERENCES

[1] T. Kozai, "Resource use efficiency of closed plant production system with artificial light: Concept, estimation and application to plant factory," *Proc. Jpn. Acad., Ser. B*, vol. 89, no. 10, pp. 447–461, 2013.

[2] Z. Tian, W. Ma, Q. Yang, and F. Duan, "Application status and challenges of machine vision in plant factory—A review," *Inf. Process. Agricult.*, vol. 9, no. 2, pp. 195–211, Jun. 2022.

[3] N. Chinsuwan and T. Radpukdee, "Mathematical modeling of a plant factory system for optimal fuzzy logic control," in *Proc. IEEE Int. Conf. Autom. Control Intell. Syst. (I2CACIS)*, Jun. 2020, pp. 195–200.

[4] S. Sabudin, M. E. Zulkarnaen, A. N. Mohammed, and M. F. B. M. Batcha, "Numerical investigation of temperature distribution in a container-type plant factory," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 28, no. 2, pp. 90–101, Oct. 2022.

[5] N. Lu, E. L. Bernardo, C. Tippayadarapanich, M. Takagaki, N. Kagawa, and W. Yamori, "Growth and accumulation of secondary metabolites in perilla as affected by photosynthetic photon flux density and electrical conductivity of the nutrient solution," *Frontiers Plant Sci.*, vol. 8, p. 708, May 2017.

[6] J. Song, H. Huang, Y. Hao, S. Song, Y. Zhang, W. Su, and H. Liu, "Nutritional quality, mineral and antioxidant content in lettuce affected by interaction of light intensity and nutrient solution concentration," *Sci. Rep.*, vol. 10, no. 1, pp. 1–9, Feb. 2020.

[7] Y. Zhang, M. Kacira, and L. An, "A CFD study on improving air flow uniformity in indoor plant factory system," *Biosystems Eng.*, vol. 147, pp. 193–205, Jul. 2016.

[8] A. M. Noh, M. A. M. Tahir, S. Mat, and M. H. Dzulkifli, "CFD simulation of temperature and airflow inside a shipping container size plant factory for optimal lettuce production," *Supplementary*, vol. 4, no. S6, pp. 54–59, Dec. 2020.

[9] M. Liebman-Pelaez, J. Kongoletos, L. K. Norford, and C. Reinhart, "Validation of a building energy model of a hydroponic container farm and its application in urban design," *Energy Buildings*, vol. 250, Nov. 2021, Art. no. 111192.

[10] T. S. Wai, C. Chaichana, and N. Maruyama, "Energy cost analysis of growing strawberries in a controlled environment chamber," *Energy Rep.*, vol. 9, pp. 677–687, Mar. 2023.

[11] G. Pennisi, S. Blasioli, A. Cellini, L. Maia, A. Crepaldi, I. Braschi, F. Spinelli, S. Nicola, J. A. Fernandez, C. Stanghellini, L. F. M. Marcelis, F. Orsini, and G. Gianquinto, "Unraveling the role of red: Blue LED lights on resource use efficiency and nutritional properties of indoor grown sweet basil," *Frontiers Plant Sci.*, vol. 10, p. 305, Mar. 2019.

[12] J. Deng, H. Zhang, X. Zhang, Y. Zheng, J. Yuan, H. Liu, Y. Liu, B. Lei, and J. Qiu, "Ultrastable red-emitting phosphor-in-glass for superior high-power artificial plant growth LEDs," *J. Mater. Chem. C*, vol. 6, no. 7, pp. 1738–1745, 2018.

[13] H. G. Choi, B. Y. Moon, and N. J. Kang, "Effects of LED light on the production of strawberry during cultivation in a plastic greenhouse and in a growth chamber," *Scientia Horticulturae*, vol. 189, pp. 22–31, Jun. 2015.
[14] Y. Fu, H. Li, J. Yu, H. Liu, Z. Cao, N. S. Manukovsky, and H. Liu, "Interaction effects of light intensity and nitrogen concentration on growth, photosynthetic characteristics and quality of lettuce (*Lactuca sativa L.* Var. youmaicai)," *Scientia Horticulturae*, vol. 214, pp. 51–57, Jan. 2017.
[15] J. H. Kang, S. Krishnakumar, S. L. S. Atulba, B. R. Jeong, and S. J. Hwang, "Light intensity and photoperiod influence the growth and development of hydroponically grown leaf lettuce in a closed-type plant factory system," *Horticulture, Environ., Biotechnol.*, vol. 54, no. 6, pp. 501–509, Dec. 2013.
[16] S. Moriyuki and H. Fukuda, "High-throughput growth prediction for *Lactuca sativa L.* seedlings using chlorophyll fluorescence in a plant factory with artificial lighting," *Frontiers Plant Sci.*, vol. 7, p. 394, Mar. 2016.
[17] T. Xiong, C. Li, and Y. Bao, "Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China," *Neurocomputing*, vol. 275, pp. 2831–2844, Jan. 2018.
[18] Y. Ito, Y. Hosoda, H. Goto, S. Toda, and K. Takayama, "Plant growth prediction for lettuces using photosynthetic information," in *Proc. CIGR World Congr.*, 2022, pp. 1–2.
[19] A. Paturkar, G. S. Gupta, and D. Bailey, "Non-destructive and cost-effective 3D plant growth monitoring system in outdoor conditions," *Multimedia Tools Appl.*, vol. 79, nos. 47–48, pp. 34955–34971, Dec. 2020.
[20] M. J. Feldmann and A. Tabb, "Cost-effective, high-throughput phenotyping system for 3D reconstruction of fruit form," *Plant Phenome J.*, vol. 5, no. 1, Jan. 2022, Art. no. e20029.
[21] T. Moon, T. I. Ahn, and J. E. Son, "Forecasting root-zone electrical conductivity of nutrient solutions in closed-loop soilless cultures via a recurrent neural network using environmental and cultivation information," *Frontiers Plant Sci.*, vol. 9, p. 859, Jun. 2018.
[22] U. Tunalı, I. H. Tüzel, Y. Tüzel, and Y. Şenol, "Estimation of actual crop evapotranspiration using artificial neural networks in tomato grown in closed soilless culture system," *Agricult. Water Manage.*, vol. 284, Jun. 2023, Art. no. 108331.
[23] N. Ullah, J. A. Khan, L. A. Alharbi, A. Raza, W. Khan, and I. Ahmad, "An efficient approach for crops pests recognition and classification based on novel DeepPestNet deep learning model," *IEEE Access*, vol. 10, pp. 73019–73032, 2022.
[24] L. Li, S. Zhang, and B. Wang, "Plant disease detection and classification by deep learning—A review," *IEEE Access*, vol. 9, pp. 56683–56698, 2021.
[25] T. Kim, S.-H. Lee, and J.-O. Kim, "A novel shape based plant growth prediction algorithm using deep learning and spatial transformation," *IEEE Access*, vol. 10, pp. 37731–37742, 2022.
[26] Y. Meng, M. Xu, S. Yoon, Y. Jeong, and D. S. Park, "Flexible and high quality plant growth prediction with limited data," *Frontiers Plant Sci.*, vol. 13, Sep. 2022, Art. no. 989304.
[27] A. Rizkiana, A. Nugroho, N. Salma, S. Afif, R. Masithoh, L. Sutiarso, and T. Okayasu, "Plant growth prediction model for lettuce (*Lactuca sativa.*) in plant factories using artificial neural network," *Proc. IOP Conf. Ser., Earth Environ. Sci.*, vol. 733, no. 1, 2021, Art. no. 012027.
[28] C. Wang, W. Pan, X. Song, H. Yu, J. Zhu, P. Liu, and X. Li, "Predicting plant growth and development using time-series images," *Agronomy*, vol. 12, no. 9, p. 2213, Sep. 2022.
[29] L. Liu, J. Yuan, L. Gong, X. Wang, and X. Liu, "Dynamic fresh weight prediction of substrate-cultivated lettuce grown in a solar greenhouse based on phenotypic and environmental data," *Agriculture*, vol. 12, no. 11, p. 1959, Nov. 2022.
[30] P. E. H. Minchin, A. C. Richardson, K. J. Patterson, and P. J. Martin, "Prediction of final weight for actinidia chinensis 'Hort1 6A' fruit," *New Zealand J. Crop Horticultural Sci.*, vol. 31, no. 2, pp. 147–157, Jun. 2003.
[31] D.-H. Jung, S. H. Park, X. Z. Han, and H.-J. Kim, "Image processing methods for measurement of lettuce fresh weight," *J. Biosystems Eng.*, vol. 40, no. 1, pp. 89–93, Mar. 2015.
[32] J.-S. Jiang, H.-J. Kim, and W.-J. Cho, "On-the-go image processing system for spatial mapping of lettuce fresh weight in plant factory," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 130–134, 2018.
[33] Z. Lin, R. Fu, G. Ren, R. Zhong, Y. Ying, and T. Lin, "Automatic monitoring of lettuce fresh weight by multi-modal fusion based deep learning," *Frontiers Plant Sci.*, vol. 13, Aug. 2022, Art. no. 980581.
[34] M. H. Kim, E. G. Choi, G. Y. Baek, C. H. Kim, B. O. Jink, B. E. Moon, D. E. Kim, and H. T. Kim, "Lettuce growth prediction in plant factory using image processing technology," in *Proc. IFAC BioRobotics Conf.*, 2013, vol. 46, no. 4, pp. 156–159.
[35] S. Nagano, S. Moriyuki, K. Wakamori, H. Mineno, and H. Fukuda, "Leaf-movement-based growth prediction model using optical flow analysis and machine learning in plant factory," *Frontiers Plant Sci.*, vol. 10, p. 227, Mar. 2019.
[36] A. Okumoto, A. Shimabuku, J. Goran, Y. Hosoda, T. Yoneyama, H. Tomitaka, Y. Usami, and H. Goto, "Simulation service of market vegetable prices for the management of plant factories," in *Proc. CIGR World Congr.*, 2022, pp. 1–2.
[37] K. Maxwell and G. N. Johnson, "Chlorophyll fluorescence—A practical guide," *J. Experim. Botany*, vol. 51, no. 345, pp. 659–668, Apr. 2000.
[38] N. R. Baker, "Chlorophyll fluorescence: A probe of photosynthesis in vivo," *Annu. Rev. Plant Biol.*, vol. 59, no. 1, pp. 89–113, Jun. 2008.
[39] J. Zhang, M. Li, Z. Sun, H. Liu, H. Sun, and W. Yang, "Chlorophyll content detection of field maize using RGB-NIR camera," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 700–705, 2018.
[40] J. Goudriaan and J. L. Monteith, "A mathematical function for crop growth based on light interception and leaf area expansion," *Ann. Botany*, vol. 66, no. 6, pp. 695–701, Dec. 1990.
[41] J. J. Huang, C. D'Souza, and W. Zhou, "Light-time-biomass response model for predicting the growth of choy sum (*Brassica rapa* var. *parachinensis*) in soil-based LED-constructed indoor plant factory for efficient seedling production," *Frontiers Plant Sci.*, vol. 12, Jun. 2021, Art. no. 623682.
[42] R. Hu, W. Jia, H. Ling, and D. Huang, "Multiscale distance matrix for fast plant leaf recognition," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4667–4672, Nov. 2012.
[43] D. Morris, "A pyramid CNN for dense-leaves segmentation," in *Proc. 15th Conf. Comput. Robot Vis. (CRV)*, May 2018, pp. 238–245.

YUYA HOSODA (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from Osaka University, Osaka, Japan, in 2017, 2019, and 2022, respectively. He was an Assistant Professor with Toyohashi University of Technology, Aichi, Toyohashi, from 2022 to 2024. Since 2024, he has been an Assistant Professor with the Graduate School of Engineering Science, Osaka University. His research interest includes image and speech signal processing.

TOTA TADA received the B.E. degree from Toyohashi University of Technology, Aichi, Japan, in 2023, where he is currently pursuing the M.E. degree. His research interest includes algorithms and data structures.

HITOSHI GOTO received the B.E., M.E., and Ph.D. degrees from Hokkaido University, Hokkaido, Japan, in 1989, 1991, and 1993, respectively. He was a Research Associate with Tohoku University, Miyagi, Japan, from 1996 to 1998; and an Associate Professor with Toyohashi University of Technology, Toyohashi, Japan, from 1998 to 2020. Since 2020, he has been a Professor with the Information and Media Center, Toyohashi University of Technology. His research interests include computational chemistry and cheminformatics for organic compounds and materials.