TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Wen, J., Abeel, T., & de Weerdt, M. (2026). Data-Driven Precision Harvest Planning via Visual Growth Forecasting. *Applied Artificial Intelligence*, *40*(1), Article 2725177. https://doi.org/10.1080/08839514.2026.2725177

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.



Applied Artificial Intelligence
An International Journal

ISSN: 0883-9514 (Print) 1087-6545 (Online) Journal homepage: www.tandfonline.com/journals/uaai20

Data-Driven Precision Harvest Planning via Visual Growth Forecasting

Junhan Wen , Thomas Abeel & Mathijs de Weerdt

To cite this article: Junhan Wen , Thomas Abeel & Mathijs de Weerdt (2026) Data-Driven Precision Harvest Planning via Visual Growth Forecasting, Applied Artificial Intelligence, 40:1, 2725177, DOI: 10.1080/08839514.2026.2725177

To link to this article: https://doi.org/10.1080/08839514.2026.2725177

© 2026 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 27 Sep 2026.

Submit your article to this journal

Article views: 228

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uaai20

APPLIED ARTIFICIAL INTELLIGENCE
2026, VOL. 40, NO. 1, e2725177 (21 pages)
https://doi.org/10.1080/08839514.2026.2725177

Taylor & Francis
by informa

OPEN ACCESS Check for updates

Data-Driven Precision Harvest Planning via Visual Growth Forecasting

Junhan Wen, Thomas Abeel, and Mathijs de Weerdt

Delft University of Technology, Delft, The Netherlands

ABSTRACT

While Artificial Intelligence (AI) techniques have been widely applied to problem-solving, their true transformative power lies in enhancing people's practice. For instance, in greenhouse agriculture, data analytics can enhance crop monitoring and resource planning; computer vision can support detecting crops and their conditions; and machine learning methods can foster tasks such as quality assessment and yield prediction. However, these technologies are often deployed only in isolated stages of cultivation, highlighting the absence of an integrated pipeline that connects them to what is desired in real-world practices. To bridge this gap, we propose CROPI – Computer vision for Resource Optimization and Precision harvest using Image data. CROPI is developed with four core functionalities: i) we propose and justify using hue as an objective redefinition of fruit ripeness; ii) we parameterize the ripeness development over time by analyzing growth monitoring images; iii) we develop predictive models to forecast the growth of new fruits; and iv) we optimize fruit-level harvest plans by aligning predicted growths with sequential demands. We demonstrate this approach using strawberries as the use case. By integrating insights from data mining and machine learning into harvest planning, CROPI lays the foundation for next-generation precision agriculture, ensuring efficiency, sustainability, and improved yield management.

Introduction

Effective and efficient planning in real-world scenarios often calls for a combination of diverse, interdisciplinary techniques to extract meaningful structure and insights from data in the wild. Such data are typically multimodal, noisy, and sometimes scarce, demanding the methodical integration of complementary methods to bridge the gap between heterogeneous inputs and abstract, behavior-level targets. For example, when optimization algorithms are used to guide future plans, their utility depends heavily on accurate representations of the current state. This, in turn, necessitates the use of artificial intelligence (AI) to process unstructured inputs into reliable, real-world assessments and forecasts. However, many existing technologies focus narrowly on isolated subtasks, resulting in a lack of integrated, end-to-end workflows. Addressing this gap requires a paradigm that unifies data-driven methods to orchestrate insights from diverse sources in a coherent and task-oriented manner.

A typical field that is currently being reshaped by information technology and AI is agriculture, leading to the call for "precision agriculture," where in-field data are utilized to perform actionable crop-management decisions (Das et al. 2015). Machine learning models are widely applied to support tasks such as yield prediction and crop counting (Bali and Singla 2021; Elhariri, El-Bendary, and Mahmoud Saleh 2023; Hassanjani Roshan, Kazemitabar, and Kheradmandian 2022; Mavridou et al. 2019; Pattnaik, Shrivastava, and Parvathi 2020). However, most existing solutions remain confined to narrow, standalone applications: for example, serving as specialized computer vision models or as systems that merely mimic human judgment. Developed in isolation, they are difficult to integrate into the broader decision-making pipelines growers require, whilst what is needed instead is integrated, interdisciplinary analytics that provide objective, actionable insights. Hence, we introduce CROPI, an interdisciplinary, end-to-end workflow that combines machine learning, computer vision, and algorithmic components into a seamless framework for a realistic challenge: using accessible in-field data to determine optimal harvesting strategies,

CONTACT Mathijs de Weerdt m.m.deweerdt@tudelft.nl Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Van Mourik Broekmanweg 6, Delft 2628 XE, The Netherlands

© 2026 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

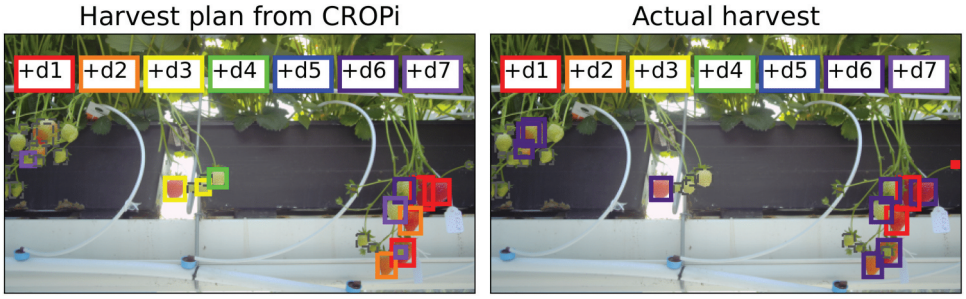


Figure 1. Composite photograph comparing CROPi harvest plan and actual harvest in a week. The left image shows a plan from CROPi under a daily demand of three fruits; the right shows the greenhouse's actual practice during the same period. In both images, "+d∗" indicates that the harvest is planned to be ∗ days later, or had occurred ∗ days from the time the image was taken.

which can assist both immediate and future decisions. An illustrative outcome of CROPI, optimizing a daily demand of three fruits across all 32 fruits observable in the field of view, is presented in Figure 1 and demonstrates its potential to support practical, data-driven harvesting decisions.

To contextualize the environment in which CROPI operates, we turn to the widely applied market model, which is also called "contract farming." Contract farming is a marketing arrangement in which growers sell most of their agricultural products to downstream buyers – typically large wholesalers with sufficient scale to exert bargaining power – with terms regarding supply quantity, quality, delivery schedules, and pricing. These pre-arranged conditions create a stable and predictable demand environment for growers, an increasingly attractive feature in modern agricultural markets, particularly for perishable products such as soft fruits (Hueth et al. 1999; Otsuka, Nakano, and Takahashi 2016). Yet this contractual stability coexists with structural asymmetries: many growers involved in these arrangements are of a small scale and operate independently, which can place them at a disadvantaged position when negotiating with larger buyers who hold greater market power (Bellemare and Bloem 2018; Hueth et al. 1999). Moreover, the consequences of breaching such contracts can be severe, while production risks are borne primarily by the agricultural industry (Bellemare and Bloem 2018; Otsuka, Nakano, and Takahashi 2016). Consequently, this forces the growers to continuously adjust decisions in their daily harvest practices, e.g. re-deciding the minimum maturity of fruits to be harvested, to meet the contracted demands without exceeding them by much. As of now, with limited tools to quantify maturity and model growth, decisions largely depend on the growers' expertise. Managing these complexities requires more advanced, data-driven approaches that can account for both analyzing growth dynamics and optimizing harvest timing.

To this end, we develop CROPI: an end-to-end workflow that leverages temporal, in-field data to optimize sequential harvest decisions. Combining computer vision, machine learning, and optimization, CROPI supports precision agriculture by aligning growth monitoring with downstream demand: in the first place, we propose using hue values as an objective alternative to eliminate inconsistencies in subjective quality assessment, together with color constancy algorithms. Second, we quantify the growth process using parametric functions (Richards 1959); while autoregressive models avoid potential curve-fitting errors, we find parametric fitting offers high accuracy with significantly lower computational cost. Third, machine learning models leverage these records to forecast future growth, enabling proactive greenhouse management. Finally, an assignment optimization algorithm integrates these predictions with demand to streamline sequential harvest scheduling.

To the best of our knowledge, no existing approach integrates the analysis and predictions of biological development with downstream demands to optimize the harvest, thereby enabling planning at the individual-fruit level. CROPI bridges this gap by mining key information from temporal monitoring data (which are mostly images), using this for forecasting daily production, and matching this to daily demands. Specifically, it converts time-lapse images into fruit-level time series that support predicting the maturity levels during growth, and disaggregates higher-level harvest demands (e.g., in kilograms) into per-fruit targets per day, so as to enable both sides of the optimization problem to be expressed at the fruit level.

Through this alignment, CROPI enables integrated, fruit-level optimization using data accessible in the field, which further paves the way for the broader adoption of more autonomous, data-driven farming techniques.

This paper begins by formally defining the problem (Section "Problem formulation"), followed by an introduction to the conceptual workflow of CROPI and the methodologies underlying each of its components (Section "Methodology"). This workflow is demonstrated using realistic strawberry growth records from greenhouse practices (Section "Experiments and demonstrations"). Strawberries are selected as the use case because their growth stages are typically reflected by changes in hue, and because open-source image datasets are available that enable individual-fruit–level monitoring (Wen, Verschoor, and Feng 2024).

Overall, the contributions of this paper lie in:

- a novel, end-to-end AI workflow that integrates multi-modal greenhouse monitoring data in a unified framework, enabling more precise and adaptive decision-making in precision agriculture;
- a vision-based method for objectively assessing fruit maturity in the field, enabling further parametric modeling approaches that support crop development predictions;
- an optimization strategy that jointly considers growth dynamics and market demand to optimize harvest decisions, providing a computationally efficient method for sequential scheduling.

Background and related work

Current practice with vision data

The development stages of many crops are defined by the change in appearance during the biological development of the plants. For fruits that show an evident color change among stages, color cards are designed to improve the consistency for people to assess the maturity level (Ayala-Silva et al. 2005; CTIFL). Hence, image analysis and machine learning are widely adapted to predict ripeness (Mavridou et al. 2019). The intuitive development of models, predicting human labels from images, can also be easily plugged with developed object detection frameworks to perform on scale. Nonetheless, in order to lift their efficiency, workers may rely more on their experience than the color cards in practice. While most existing works focused on classifying the maturity stages that are assessed by people, they essentially introduce subjectivity to the objective assessments. Thus, in this paper work, we propose redefining ripeness labeling by quantitative color analysis to achieve true objectiveness in the modeling. We believe this leads to more consistent results and better defines a continuous space of ripeness rather than manual categorization with vague and potentially overlapping boundaries.

To use color as the primary characteristic for defining ripeness, it is crucial to minimize illumination variations caused by factors such as the time of image capture and changes in viewing angles (Sala et al. 2020). Preliminary research on identifying fruit ripeness using image data was usually conducted with laboratory settings to limit the influences (Jha, Chopra, and Kingsly 2007; Lee, Archibald, and Xiong 2011). In fact, computer vision techniques such as color correction algorithms (Gijsenij and Gevers 2011; Goatman et al. 2003) and generative algorithms (Cai et al. 2024; Z. Li and Snavely 2018) are practical for conducting robust and adaptive corrections. We believe these promising candidates can facilitate the use of in-field data, which leads to higher applicability of the entire machine learning framework.

Formulation of growth functions

In agricultural research, parametric functions have been used to model crop growth, including overall development, future, and even seasonal production (we use the term "growth function" from here on). The sigmoid function, or the more generalized form, Richard's function (Richards 1959), is a widely used mathematical model to parameterize the growth of many crops (Gerhard and Moltchanova 2022; Jannatizadeh et al. 2023). In research where the growth models were considered to optimize harvests, researchers took fixed coefficients from statistics of species (Harel, Edan, and Perlman 2022). However, when other influencing factors, such as the growing environment and cultivation treatments, are omitted, the growth models are not precise enough to provide realistic predictions applicable to actual harvests (Hopf et al. 2022).

A major reason for this limitation is the lack of comprehensive analyses that account for variations in time, location, and treatment conditions. One promising way to address this gap is through the mining of crop growth monitoring data, which, as previously discussed, provides a directly accessible source of information for capturing these variations. In this context, computer vision techniques play a key role by enabling objective analysis and helping extract consistent, generalizable insights from the data.

Prediction of future growth

Looking from the dimension of time, growth functions are time-ordered data, typically shaped by species-specific patterns but influenced by environmental conditions. Forecasting growth, then, becomes a task of completing the function based on partial observations, i.e., predicting future development from what has been observed so far. This forecasting task can be framed as a next-token prediction problem, where the model continues the observed sequence by mimicking its underlying process. This approach, common in natural language processing (Achiam et al. 2023), has also proven effective in traditional time-series forecasting (Qiu Nate Gruver, Finzi, and Gordon Wilson 2023; Zhang et al. 2024).

Alternatively, since the growths tend to follow a similar pattern, k-nearest neighbor (k-NN) and clustering methods can help advance analysis by capturing temporal dependencies and complex trends through distance-based comparisons (Ahmed et al. 2023). Besides traditional measures like Euclidean distance, specialized metrics such as autocorrelation and Dynamic Time Warping (DTW) (Sakoe and Chiba 1978) could be more practical to be considered. When the data contains noise or perturbations, distances like softDTW (Cuturi and Blondel 2017) and LB Keogh lower bound (Keogh and Ratanamahatana 2005) are preferred, as they offer greater robustness by accommodating temporal mis-alignments and outliers. With these metrics, the aforementioned machine learning methods could work more effectively.

Harvest and supply optimization

Existing research on agri-food supply optimization has typically addressed broader, more general problems, on which researchers often employed linear programming or dynamic programming models to incorporate uncertainties and constraints (Omar and Rene Villalobos. 2009; Taşkıner and Bilgen 2021). The implementation of such methods tailored the solutions to focus on decision-making at the level of management units or batches (Catalá et al. 2016; Gómez-Lagos et al. 2021). However, this approach remains distant from the practical realities faced by harvest robots. Alternatively, in solutions for other practices, when the optimization is framed as an assignment problem and subjected to proper feasibility checks (Cacchiani, Caprara, and Toth 2010; T. Li, Li, and Qian 2016), the Hungarian algorithm can achieve optimized planning within polynomial time (Kuhn 1955; Lavanya and Santhi 2021).

Problem formulation

Context and definition

In this paper, we investigate a real-world use case that integrates analytics of fruit (growth) development with the optimization of harvest decisions. For non-climacteric fruits like berries, which do not keep ripening after picking, their maturity at harvest directly determines selling quality. Therefore, CROPI primarily focuses on optimizing the timing of harvest by aligning fruit growth stages with delivery requirements.

The problem involves two fundamental sources: on one side, cameras in greenhouses monitor fruit growth through videos or time-lapse images, enabling us to reconstruct past development and predict future ripeness by associating changes with frame timestamps. On the other side, delivery contracts provide known requirements, specifying quantities and deadlines in advance. Hence, we model the problem as an assignment optimization: allocating harvest times to fruits to best meet quality and delivery targets.

Formulation

Overall, the problem is formally defined as:

- **Given**:

- The growth monitoring of fruits $j \in \{1, 2, \ldots, N\}$, in the form of video frames or image series.
- The demand from downstream $D_t, t \in \{1, 2, \ldots, T\}$, which are sequential and are fixed from contracts.
- Other related records, e.g. weather in the past.

- **Find**:

- The optimal plan, defined as a set of binary actions x_{jt}("harvest or not") over time t for each fruit j.

The corresponding optimization problem is formulated as:

- **Objective**:
- Maximize the profitability of selling by planning x_{jt}:

$$\max\left(\Sigma_{t=1}^{T}\Sigma_{j=1}^{N} x_{jt} \cdot V_{jt}\right) \tag{1}$$

- Where the profit V_{jt} is assessed by scoring the growth of fruit j at the proposed harvest time t.
- **Constraints**
- As a natural constraint, one action per fruit, i.e. $\Sigma_{t=1}^{T} x_{jt} \leq 1, \quad x_{jt} \in [0, 1]$ for $j \in \{1, 2, \ldots, N\}$.
- Each growth is modeled as an independent time-dependent function $G_j(t)$, i.e. harvesting j will not significantly impact others' $G_{_(!j)}(t)$ of the plant.

Data source

In the demonstration of this paper, we primarily utilized public image datasets to evaluate the approach under a spectrum of demands. Our inputs cover the following categories:

Greenhouse monitoring

"The Growing Strawberries" Dataset (GSD) (Wen, Verschoor, and Feng 2024) collected on-site time-lapse images during cultivations in two years. Example raw images are depicted in the first row of Figure 6. Its annotations for a multiple object tracking (MOT) task identify the growth of individual strawberries, from which we can curate datasets for growth monitoring and forecasting in the following demonstration.

In addition, we used the timestamps of the monitoring images to align the cultivation record with on-site climate computer measurements (Wen et al. 2023) and close-site weather data (Hersbach et al. 2023), so as to provide more input features for the prediction task.

Future demands

Due to the lack of relevant public data, we designed a set of typical comparative scenarios to simulate the demands required for our analysis. These scenarios represented cases in which demand could be higher than, close to, or lower than projected future yields. The synthetic data was crafted to reflect realistic conditions and generate meaningful insights for the study.

Fruit maturity assessment

In addition, we discuss the outcomes of our proposed objective ripeness assessment in Ripeness evaluation, and compare them with descriptive assessments or discrete scores assigned using typical datasets provided by (Elhariri, El-Bendary, and Mahmoud Saleh 2023) and (Wen et al. 2023) respectively.

Methodology

Figure 2 depicts the workflow of our proposed solution. The solution utilizes accessible on-site data from real-world greenhouses to optimize the sequential harvest decisions based on predictive growth models of fruits, so as to support the decision-making of how much and which fruits to harvest for the best possible match to the demand over time. The methodologies of the four main functionalities, as identified by the diamond shapes, are introduced in the following subsections, and demonstrated in Ripeness evaluation to 5.4 respectively.

Processing the images

The right subplot of Figure 2 illustrates the main and branched streamlines of the image processing workflow. Given videos or image series capturing crop growth, multiple object tracking (MOT) algorithms can identify individual strawberries and therefore reveal the individual growth pattern of every fruit. To reduce noise in the further processing, segmentation models, or taking a subarea in the detected bounding box as an approximating alternative, can be applied on top of the MOT results. In the meantime, since many cultivation environments use only limited lighting control, color consistency correction methods such as the Gray World (Goatman et al. 2003), intrinsic color decomposition (Z. Li and Snavely 2018), or style transfer models (Cai et al. 2024) can also be added to the main workflow.

We chose strawberries as an example because fruit color is a widely accepted indicator of maturity (Made Defy Janurianti 2021), making it a simple yet representative case for demonstrating CROPi. Traditionally, ripeness is assessed by comparing fruits to reference color cards or personal expertise. While computer vision offers more consistent and objective evaluations, current methods often rely on subjective human annotations to train models from scratch or fine-tuning from object detectors like YOLO or R-CNN. Since biological development is continuous and lacks clear stage boundaries, we propose using the fruit skin's hue to eliminate human error and ensure true objectivity. In the specific case of strawberries, we used the A* value from the CIELAB color space (Schanda 1996), which reflects their natural growth in a straightforward manner – from green on the negative axis to red on the positive axis. The A* values collected over time form *(timestamp, color)* pairs that serve as the raw data for growth modeling.

Modeling the growths

We considered two ways of developing growth models: (1) a traditional method that fits development using a logistic growth function, referred to as the "parametric growth function," and (2) a deep learning approach inspired by natural language processing (NLP), which models growth as a time-based sequence using an autoregressive next-token prediction (NTP) strategy. Given its roots in NLP techniques, we refer to the second approach as the "growth language" model.

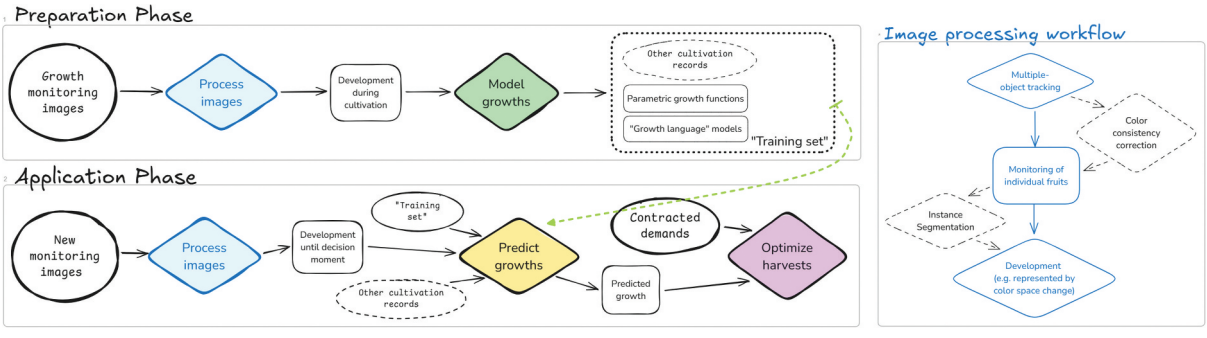


Figure 2. Pipeline of the CROPi framework, divided into two main sections: the preparation phase and the application phase. The core functional components, represented by the four color-coded diamond shapes, are benchmarked and discussed respectively in subsections of Section "Methodology" to Section "Experiments and demonstrations". Specific steps of the first stage, "process images," are provided within the additional section on the right.

Fit to parametric growth functions

Both Richard's curve and the logistic function are suitable for modeling the typical smooth-rapid-smooth growth pattern, as illustrated in Figure 6. Between them, the fewer coefficients for the logistic function, as shown in Equation (2), make it more robust and efficient for fitting noisy real-world data (Jannatizadeh et al. 2023) and thus make it more applicable for our case. Example growth functions are shown as dashed lines in Figure 7. Its alignment with natural biological growth trends also allows us to identify potential detection or tracking errors by comparing the fitted parametric curve to the raw A* value trajectories. Such a comparison also highlights potential errors in GSD annotations, helping us refine the dataset before training the predictive models.

$$f(t) = \frac{A}{1 + e^{-B(t-t_0)}} + C \tag{2}$$

Note that since MOT is not the main task of CROPI, we used the revised annotations in the Experiments and demonstrations section. Moreover, existing benchmarks from (Wen, Verschoor, and Feng 2024) indicate that the primary bottleneck for state-of-the-art MOT algorithms on GSD lies in the track ID association stage, whereas the object detection performance remains on par with the performance on larger and more widely used datasets. Consequently, relying on ground-truth annotations enables us to obtain more complete and reliable growth tracks, which in turn provides stronger supervision for training the subsequent prediction models. For future applications on new monitoring data, we can still employ real-time MOT algorithms to keep track of each new fruit's growth.

Train growth language models

Beyond parametric function models, deep learning offers a more flexible way to capture growth patterns. Since input data lengths vary in realistic applications, NTP provides a practical mechanism for fully leveraging all available observations. In parallel, the increasing availability of large language models (LLMs) has made transfer learning particularly appealing for domains with limited datasets, such as agriculture. Accordingly, in developing CROPI, we explored both training NTP models from scratch and from fine-tuning open-source LLMs to learn the growth patterns of the crops.

In experiments in Growth prediction, we applied *nanoGPT* (Karpathy 2022) to build our *growth language* models. We selected nanoGPT because its computational requirements remain manageable on a single NVIDIA A40 GPU, and its modular architecture integrates effectively within our multi-stage workflow without requiring substantial engineering effort. As the computational demand is still very high, we limited each model's training time to a maximum of 36 hours.

Predicting the growths

For growth prediction, the goal is to estimate future development over a fixed time range based on the ripeness states observed so far (i.e., MOT results up to that moment) for strawberries in a given frame. To this end, we used the GSD dataset (Wen, Verschoor, and Feng 2024), which provides both image series and MOT annotations from 2021 and 2022, and enables two train–test splits: (i) training on one camera's 2021 series and testing on another camera's 2021 images, and (ii) training on all 2021 images and testing on a single camera's 2022 series.

In addition to the image data, we incorporated near-site weather. For 2021 we also have on-site micro-climate data available (Wen et al. 2023), but did not include this because it would limit the possibility of evaluating cross-season generalization.

When using the *growth language* model, we treated each series as independent and added separating tokens at the end. However, a fixed forecast length can cause the model to predict a separating token prematurely, triggering a "new" series to be predicted without any observation. To address this, we replace all predictions after the first separating token with the last valid token before it, until the forecast period ends. This approach reflects the idea that maturity cannot decrease, and the extended stay in the mature stage can be handled later using the scoring function described in Optimizing the harvest plans.

When fitting the *parametric growth function*, we first filtered by min. length, max. initial, and min. ending to ensure the task uses relatively more complete growth monitoring records throughout the seasons.

These, together with relevant climate data from near-site sources, formed the "knowledge" of the model, which is also considered as the "training set."

Since the training and testing sets cover different periods but share the same hourly resolution, we developed a two-stage algorithm to estimate growth function coefficients: In the first stage, we identified growth functions from the training set whose inflection points (t_0 in Equation (2)) fall within a narrow time window. The primary inflection point T'_0 of the testing set was determined by aggregating votes from growth observations that have passed the lag phase (Fang et al. 2022), i.e. which are characterized by a clear transition from green to red, by comparing pattern similarities with historical growth observations or climate data in the training set. In the second stage, after narrowing down the candidate growth functions based on T'_0, we estimated the fruit-specific inflection dates $T'_{j,0}$ and the associated coefficients A, B, and C of Equation (2). These parameters were obtained by reapplying the similarity measures to identify the most compatible growth function for each fruit j. To assess these similarities, clustering, autocorrelation (autocorr.), and sliding-window k-NN (s.k-NN; as illustrated in Figure 3) were used for growth data, while clustering, 2D-autocorr., and k-NN were applied to 2D data of historical climate. S.k-NN with the same metric was used in the second stage if the first stage used a clustering method.

Optimizing the harvest plans

The ultimate task of CROPI is to maximize the quality of supplies through planning the sequence of harvesting actions. We model this as an assignment optimization problem, where each fruit is assigned an optimal harvest time to ensure the best possible outcome. For non-climacteric fruits such as strawberries, which stop growing once harvested, the problem is equivalent to maximizing the overall quality at each fruit's harvest moment. The reward of harvesting a fruit j at time t, denoted as $\mathbf{V}_{jt}$, can be assessed by the growth stage, $G_j(t)$. In Figure 4, we illustrate a few scoring options – such as binary, fuzzy logic, and fuzzy logic with timing penalties – which serve as examples but are by no means exhaustive.

We treated the harvesting decisions, denoted as a binary variable x_{jt}, as optimizing the assignment between fruit growths and harvest time slots. Hence, the fundamental form of the objective function, which is commonly referred to as the cost matrix in assignment problems, is defined as follows:

$$\max \quad \Sigma_{t=1}^{T}\Sigma_{j=1}^{N} x_{jt} \cdot V_{jt}$$

$$s.t. \quad \Sigma_{t=1}^{T} x_{jt} \leq 1, x_{jt} \in [0,1] \text{for} j \in \{1,2,\ldots,N\} \tag{3}$$

Building on this formulation, we discretized the greenhouse-level demand based on visible plant area, resulting in integer-valued demands for each time slot, i.e., $D_t \in N, \quad t \in \{1,2,\ldots,T\}$. This setup enables the use of the Hungarian method to perform the optimization, which also guarantees that each element in the smaller set is matched to one in the larger set (Kuhn 1955). As a result, if supply exceeds demand, some fruits will not be scheduled for harvest; if demand is too high, even the greenest fruits may be assigned. To address these impractical scenarios, we introduce two additional options: an *oversupply allowance* D'_t,

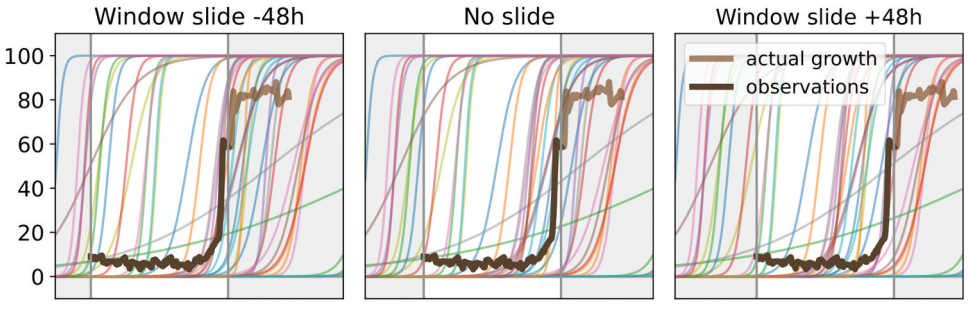


Figure 3. Demonstration of the s.K-NN method. The observation is shown in dark brown, with lighter shades indicating the ground-truth future. Thinner lines are growth functions from the training set. For each observation, the k functions with the highest similarity and the corresponding window sliding pointers are noted to estimate $T'_{j,0}$.

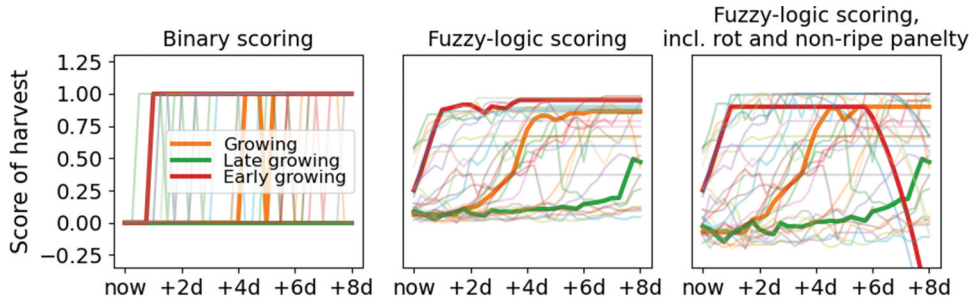


Figure 4. Three example scoring methods for harvests at different future moments, applied to predictions from our nanoGPT model. Each panel shows a different scoring method applied to the same data. We highlighted the score change from three typical observation patterns, as noted in the legend.

representing an alternative supply venue where the demand and/or price might be (δ-)discounted; and a *postponed harvest* by placing a dummy demand variable D_∞. The corresponding decisions are denoted as y_{jt} and z_j, respectively. With these options included, the objective function is updated accordingly:

$$\begin{aligned} \max \quad & \Sigma_{t=1}^{T}\Sigma_{j=1}^{N}(x_{jt}\cdot V_{jt} + y_{jt}\cdot\delta\cdot V'_{jt}) + \Sigma_{j=1}^{N} z_j \cdot V_\infty \\ s.t. \quad & \Sigma_{t=1}^{T}(x_{jt} + y_{jt}) + z_j = 1,\ \forall j \in \{1,2,\ldots,N\} \end{aligned} \tag{4}$$

Hence, when $x_{jt} = 1$ or $y_{jt} = 1$ in the solution, the optimal harvest time for that fruit j is the corresponding t. If $z_j = 1$, the fruit is excluded from this given harvesting period t.

Since no established metrics exist for evaluating fruit-level harvest activities, we assess the outcome plans by estimating the "reward," i.e. $V_j(t)$ as in Equation (4), assuming a constant daily demand per fruit. This assumption reflects both the unavailability of real-world grower-specified demand data and the fact that a substantial portion of sales is predetermined through contractual agreements, making daily demand effectively constant within our discretized framework. In the section "Harvesting optimization" (Experiments section), we compare the plans optimized using predicted fruit growth with three benchmarks: (i) the actual harvest, derived from MOT annotations in *GSD*; (ii) a simple "harvest all mature fruit" policy, applied directly to the recorded growth data; and (iii) optimized plans using *parametric growth functions* fitted to those records ("complete simulation" in Table 3). To ensure a fair comparison, we applied a fuzzy-logic scoring system that accounts for penalties, as later visualized in Figure 4, under the simulated parametric growth setting to calculate $V_j(t)$.

Experiments and demonstrations

This section demonstrates the methodologies employed to execute the four principal functions of the solution pipeline, corresponding to the color-coded components in Figure 2.

Ripeness evaluation

To illustrate the necessity of using color as an objective measure to assess ripeness, we first compared the exact color values from the fruit with the maturity labels provided by the datasets, which are annotated by agricultural experts (Elhariri, El-Bendary, and Mahmoud Saleh 2023; Wen et al. 2023). For computational efficiency and to limit the influence of background pixels, we extracted the center 50% of each detection bounding box (bbox) and computed the average A* value over this region. We compared these assessments with the original ripeness labels provided by the two datasets: *StrawberryDS* from Elhariri et al. (Elhariri, El-Bendary, and Mahmoud Saleh 2023) that uses six descriptions to categorize fruits from early growth until peak ripeness, and the quality assessment collection from Wen et al. (Wen et al. 2023) that uses a discrete grade from 1 to 10 to identify the maturity levels. Note that the latter dataset only consists of harvested strawberries due to other measuring purposes, resulting in the labels not being distributed over all values.

Figure 5 illustrates a trend where A* values generally increase with higher ripeness levels. However, some overlap exists between distributions, especially when the categories are too fine. These overlaps likely stem

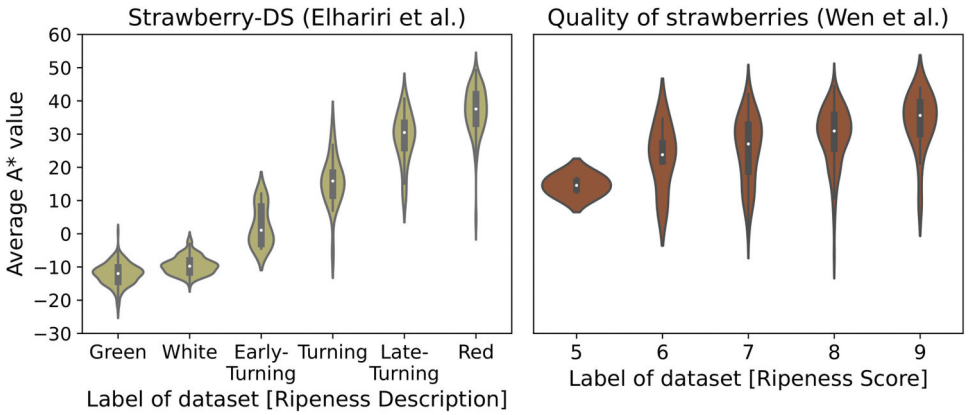


Figure 5. Distribution of A* values of fruits in images under different ripeness labels. The left comparison uses the training set from Elhariri et al. (Elhariri, El-Bendary, and Mahmoud Saleh 2023). The right plot uses the full dataset from Wen et al. (Wen et al. 2023). The original annotation of each dataset is indicated on the x-axis of the corresponding figure.

from the subjectivity of human labeling or occlusion by other objects in view, with subjectivity being the primary factor. This is particularly significant because such ambiguity is even more evident in (Wen et al. 2023), where the classification boundaries are vague due to class refinement. Hence, the inherent differences between human and machine labeling efforts further underscore the need for objective metrics.

This experiment validates using hue (or A* in this case specifically) as the primary reference rather than just a supplementary aid, to reduce subjectivity and improve consistency in ripeness evaluations. This approach not only supports more transparent and data-driven supply negotiations but also lays the foundation for building a quantitative, continuous growth model, as further demonstrated in the following functionalities of CROPI.

Growth function formulation

When we applied an objective ripeness evaluation as presented in Ripeness evaluation to process growth monitoring data, we could record the biological development as a time-based function. In Figure 7, we demonstrated this methodology using the growth monitoring images and MOT annotations of GSD (Wen, Verschoor, and Feng 2024).

As shown in Figure 7, the A* value generally increases over the stages of biological development, though at varying rates. However, notable noise causes fluctuations, primarily due to changes in illumination conditions during the monitoring period. Among these, the most significant variations can be eliminated by filtering image data using a brightness threshold, as illustrated by the half-transparent lines in Figure 7.

To further mitigate the impacts of the varying illumination, we tested a few color correction methods on the image data, as demonstrated in Figure 6. The growth functions after selected pre-processing are shown as other colored lines in Figure 7. A comparison of Figures 6 and 7 shows that intrinsic image decomposition reduces temporal fluctuations, but at the cost of diminished image contrast and a narrower range of hue values. Generative models used for style transfer also improved color consistency from a perceptual standpoint; however, they sometimes introduced unpredictable sharp values, which adversely affected further processes. In contrast, although the Gray World algorithm offered a more modest adjustment, its deterministic nature ensures that it behaves consistently and avoids the instabilities of generative approaches. Thus, we used its output in subsequent demonstrations.

Growth prediction

Given the available crop observations up to the decision-making point, we frame growth prediction as a completion task – estimating the full developmental trajectory from incomplete data. To this end, we explored two approaches: 1) training a *growth language* model to perform autoregressive next-token

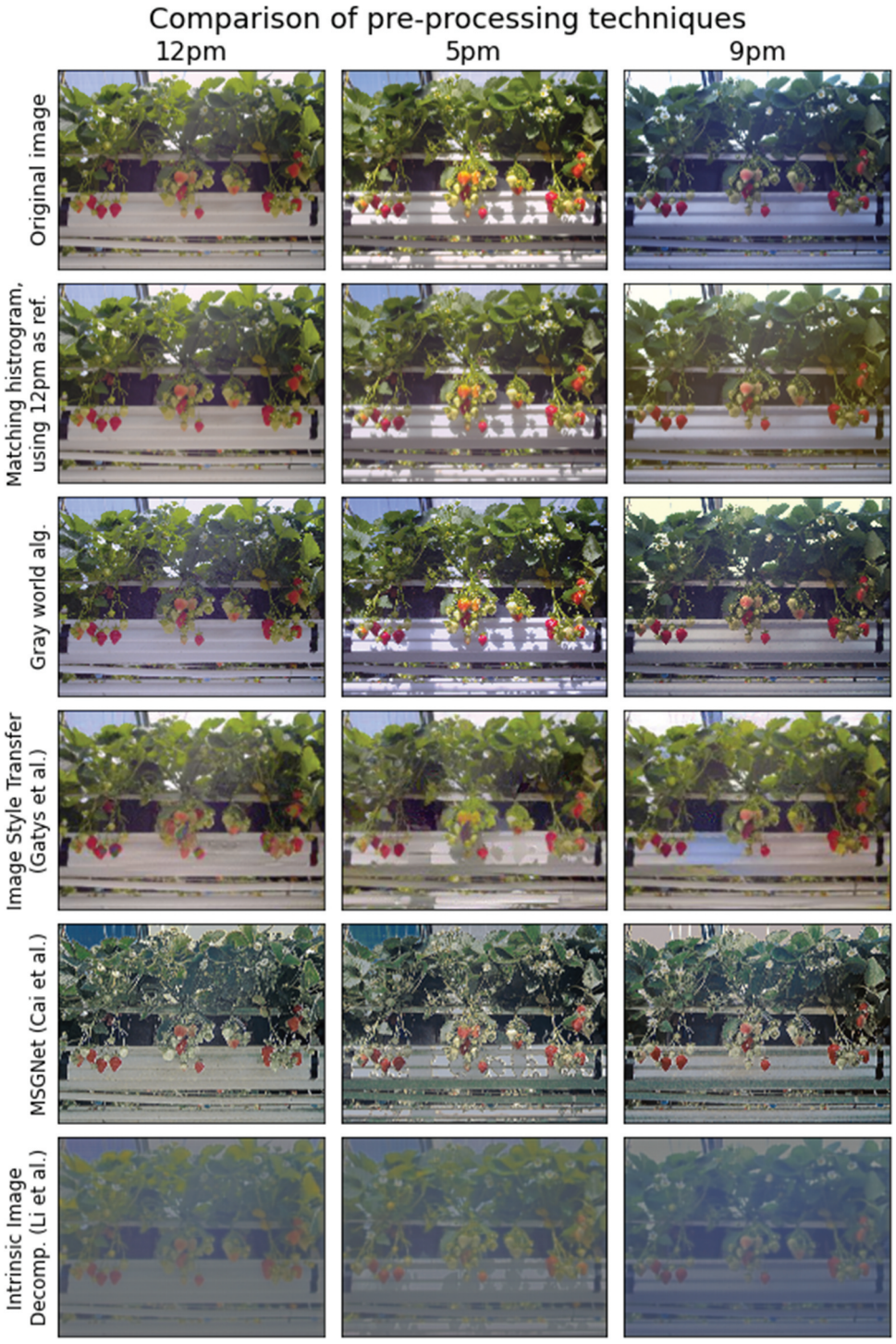


Figure 6. Illustration of five color consistency correction methods, using sample images collected at 12 pm, 5 pm, and 9 pm. Image Style Transfer (Gatys et al. 2016).

prediction (NTP), and 2) fitting the *parametric growth function* from Equation (2) by seeking similar samples from observed patterns in the training set.

Table 1 and 2 present the prediction results for the two train-test splits described in Predicting the growths. Overall, NTP-based methods demonstrate stronger performance for this task. Notably, since most models trained from scratch did not fully converge within the budget computational time, the accuracy achieved by fine-tuning a pretrained GPT-2 model stands out more clearly. This advantage is particularly evident under cross-entropy loss, which favors sequence modeling tasks but is less suited for continuous regression.

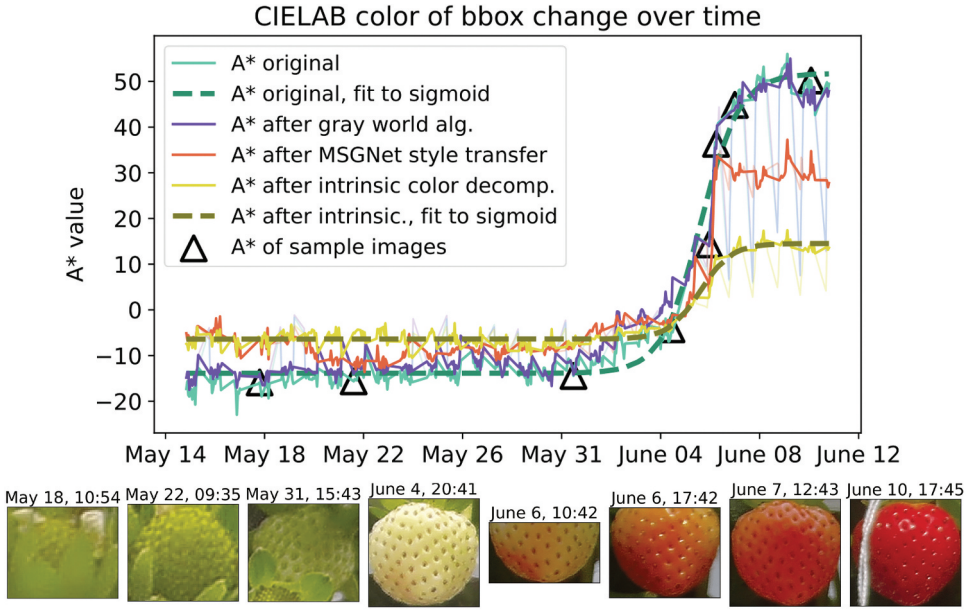


Figure 7. Example growth function (top) and observations (bottom) for a strawberry. A* Values are shown across the original images and different color-correction methods. Dashed lines represent sigmoid curves fitted to the A* values. Darker lines indicate data from images with brightness above 50. Triangles mark the data points of the below observations.

Table 1. Comparison of selected growth prediction methods, measured on all visible strawberries on a random frame from 2021 as in the harvesting optimization task. The table is split into four sections, based on method and data source. Performances are described by common time-series analysis metrics shownin the header. Note that for two-stage methods, the same metric as in the second column was used for both stages. Superscripts link methods to their optimization performance in Table 3.

Method	Metric	L2 ↓	DTW ↓	AC ↑	Time [s]
PERFECT SIGMOID FIT		0.19	1.73	0.13	–
TRAINING A NEXT-TOKEN PREDICTION MODEL					
Full training	L2 Norm	0.19	2.84	0.09	(max 36 h)
Full training	CrossEntropy	0.33	3.68	0.12	(max 36 h)
Finetune GPT-2✪	L2✪	**0.17**	2.30	0.10	(max 36 h)
Finetune GPT-2	CrossEntropy	0.18	2.01	0.08	(max 36 h)
CONSIDERING SHAPE SIMILARITY OF GROWTH FUNCTIONS					
autocorr.	autocorr.	0.75	8.74	0.30	9
s.k-NN	L2 Norm	0.22	2.81	0.10	4
s.k-NN	DTW	0.23	3.05	0.15	157
s.k-NN☆	LB_Keogh☆	0.21	**2.29**	0.13	232
K-means	L2 Norm	0.26	2.65	0.08	28
K-means	DTW	0.30	3.48	0.06	31114
CONSIDERING SIMILARITY IN HISTORY CLIMATE					
2d-autocorr.	autocorr.	0.65	7.79	**0.36**	10
K-NN✳	cos-sim. ✳	0.25	2.78	0.00	10
K-means	L2 Norm	0.26	2.65	0.08	42
K-means	DTW	0.28	2.69	0.05	202

In contrast, the pattern-matching approaches achieved comparable accuracy to the NTP-based methods, with sliding-window k-NN (s.k-NN) attaining the lowest DTW score of 2.29, as shown in Table 1. Additionally, on the unseen 2022 data, some of the methods demonstrated better robustness than auto-regression models trained from scratch, providing a stable and reliable alternative when pre-trained weights are unavailable. Consequently, pattern-matching methods offer a lightweight, training-free option, enabling immediate inference without the extensive pre-training required by transformer models.

Furthermore, within the same methodology family, the matching based on growth patterns outperformed those relying on the similarity of the climate data. This suggests that climate similarity, particularly

Table 2. Comparison of selected growth prediction methods, trained on the full 2021 dataset and evaluated on all strawberries in a random frame from 2022. A subset of methods that performed relatively better in the 2021 experiment (Table 1) is included. Performances are described by the same metrics as in Table 1, with each row showing the combination of method and data source in the first column.

Method & Data Source	Metric	L2 ↓	DTW ↓	AC ↓
PERFECT SIGMOID FIT		0.17	1.53	0.18
PREDICTION PERFORMANCES				
Train NTP from scratch	L2 Norm	0.27	2.86	0.20
Finetune GPT-2 for NTP	L2 Norm	0.18	1.68	0.19
Finetune GPT-2 for NTP	CrossEntropy	0.19	2.02	0.19
s.k-NN on $G_j(t)$	L2 Norm	0.20	2.35	0.16
s.k-NN on $G_j(t)$	LB_Keogh	0.21	2.41	0.17
K-means on $G_j(t)$	L2 Norm	0.24	2.67	0.15
2d-autocorr. on climate record	–	0.37	3.63	0.00
k-NN on climate record	cos-sim.	0.37	3.63	0.00

when assessed via rigid metrics like k-NN, is an insufficient criterion for selecting candidate growth functions from a narrow time window in the first stage, especially when training and testing data came from different years. Meanwhile, autocorrelation (autocorr.) proved ineffective for both growth and climate matching. Since most of the growth observations follow a similar sigmoid-like pattern, varying mainly in growth velocity, simply capturing the general trend lacks the discriminative power to distinguish individual fruits. Consequently, these methods often misidentified candidate growth functions, resulting in the overly averaged, poor predictions observed in Table 2.

Even in the testing phase, a trained model would take around 2 minutes to perform the NTP task. Hence, among the methods presented, s.k-NN offered a favorable balance between performance and runtime. In addition to this, as can be observed from Figure 4, outputs from NTP often retain fluctuations caused by illumination noise in the input data. To ensure clearer visualizations and more stable results for downstream analysis, we used predictions from *s.k-NN with LB_Keogh distance* in the main demonstrations of Harvesting optimization and in Figure 1. We also evaluated how different levels of prediction accuracy influenced optimization performance with more prediction results (with superscripted markers in Table 1).

Harvesting optimization

We applied CROPI to a randomly selected timestamp – the 1000th frame from Camera RGB-2 in the 2021 split of the GSD dataset (Wen, Verschoor, and Feng 2024) – and visualized the following in Figure 8: observed growths until the decision moment, predicted growths using s.k-NN with LB_Keogh distance75, changes in harvest reward (which we simply named the "score") over predicted growth, and suggested harvest plans under varying demands. In the third subplot, dashed vertical lines indicate the time points at which the suggested plan recommends harvesting, i.e., terminating simulated growth for a non-climacteric fruit such as strawberry, assuming a daily demand of three fruits from the plants within the camera's field of view.

The final harvest plan, illustrated in Figure 1, was optimized under a daily demand of three fruits. Harvest gains were evaluated using fuzzy-logic scoring, which incorporated penalties for harvesting either rotten or unripe fruits, as illustrated in the third subplot of Figure 4. This scoring was based directly on growth predictions, represented by the A* values, which were normalized according to the observed minimum and maximum after excluding outliers. The optimization employed the expanded objective function from Equation (4), incorporating both a postponed harvest option with $V_\infty = 0$ and an oversupply option defined as $V'_{jt} = 0.5 \times V_{jt}$. Using this framework, we further examined the impact of varying daily demand levels to evaluate the robustness of the harvest plans, as shown in the final subplot of Figure 8.

Since there are in total 32 fruits in view, we consider a daily demand of three, four, or five fruits to represent cases in which total weekly demand is lower than, comparable to, or exceeds the capacity of the plant in view, i.e. the visible fruits that are subject to the optimization. As shown in Table 3, CROPI significantly improved the overall outcomes. When using the most accurate growth predictions, the optimized plan under a daily demand of three fruits achieved a harvest score of 18.75 with full demand

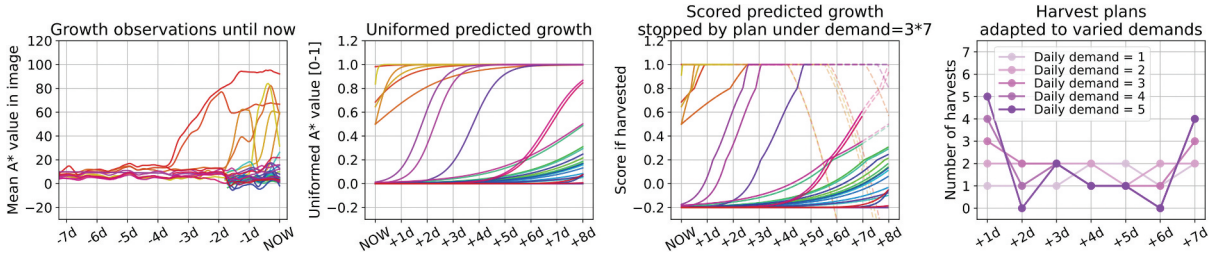


Figure 8. Four illustrative subplots show color-coded observed growth, predicted growth, scored predictions, and the resulting harvest plan, respectively. The example uses fuzzy-logic scoring with penalties, as shown in subplot 3 of Figure 4. The final subplot compares harvest plans optimized under different demand levels, using the scored predictions shown in the third subplot.

Table 3. Comparison of harvest plans under multiple daily demands using 2021 scenario predictions. Plans were evaluated by applying fuzzy-logic scoring with Equation (4). Demand fulfillment was measured with only the basic daily demands. The results with predictions, marked with the special superscripts, were from the outputs of corresponding models as in Table 1.

Daily	Harvest	Optimized	Demand
Demand	Method	Score	Fulfillment[%]
d = 2	Actual harvest	5.60	28.57
	Harvest once mature	8.22	50.00
	Opt. w/ ground-truth simu.	19.36	100.00
	Opt. w/ prediction from pred$^{\clubsuit}$	17.39	100.00
	Opt. w/ prediction from pred*	12.71	100.00
d = 3	Actual harvest	5.60	28.57
	Harvest once mature	8.22	47.62
	Opt. w/ ground-truth simu.	20.39	100.00
	Opt. w/ prediction from pred$^{\clubsuit}$	18.75	100.00
	Opt. w/ prediction from pred*	13.41	76.19
	Opt. w/ prediction from pred*	12.00	57.14
d = 4	Actual harvest	5.60	25.00
	Harvest once mature	8.22	46.43
	Opt. w/ ground-truth simu.	21.10	75.00
	Opt. w/ prediction from pred$^{\clubsuit}$	20.01	92.86
	Opt. w/ prediction from pred*	14.52	64.29
d = 5	Actual harvest	5.60	22.86
	Harvest once mature	8.22	42.86
	Opt. w/ ground-truth simu.	21.10	62.86
	Opt. w/ prediction from pred$^{\clubsuit}$	20.16	77.14
	Opt. w/ prediction from pred*	15.20	57.14

fulfillment, closely approaching the optimum obtained using ground-truth growth data. Less precise predictions*,* resulted in lower scores and smaller demand fulfillment, illustrating how errors in intermediate growth predictions propagate into suboptimal harvest decisions. In contrast, the current harvest plan based on an empirical approach and derived from GSD annotations – yielding two fruits on day 1 and eight on day 6 – performed worse overall in both harvest score and demand fulfillment, thereby demonstrating the value of optimization-based planning.

As illustrated in the final subplot of Figure 1 and Table 3, harvests with CROPI could have provided a steadier supply with more balanced harvest schedules. In specific, when demand is high, it responds by scheduling more fruits for harvest the next day ("1d"), aligning closely with observed real-world behavior; while for lower demand, it allows slight postponement to optimize quality and applies moderate overharvesting to prevent loss. Across daily demand levels of 3–5 fruits, CROPI maintained balanced harvest distributions, improved cumulative scores, and demonstrated strong potential for stabilizing supply chains through predictive, sequential decision-making.

Furthermore, the current setting of such a sparse harvesting schedule is partly due to labor constraints, as human workers cannot cover all plants efficiently. This highlights the potential of robotic harvesting, yet a simple "harvest once mature" policy remains suboptimal, particularly for maintaining a stable supply. While prediction uncertainty persists, CROPI strategically optimized harvest timing, mitigating these

challenges and demonstrating the potential of AI-driven solutions that enhance both adaptability and efficiency in precision agriculture.

Ablation studies

Beyond the demonstrations of the four main components of CROPI, we conducted two targeted ablation studies. The first study extended the analysis in Predicting the growths by examining the effect of different color normalization methods, thereby validating the design choices described in Modeling the growths. The second study assessed the impact of incorporating alternative options, i.e. oversupply and postpone harvest, as formulated in Equation (4), using both the proposed scoring and observed harvesting outcomes. These analyses motivate the current configuration and identify directions for further refinement.

Impact of color representation and normalization

The impact of color normalization and ripeness representation choices is summarized in Table 4. Using raw RGB images resulted in higher fitting and prediction errors due to residual illumination variation and sensor noise. Intrinsic image decomposition substantially reduces the fitting error ($L_2 = 1.5$) by compressing the reflectance signal into a narrow value range, as also shown in Figure 7. While this variance suppression improved curve fitting and yielded low pointwise prediction error, the required rescaling amplified small temporal deviations, leading to higher DTW distances and reduced sequence-level similarity. In contrast, Gray World normalization preserved sufficient chromatic variation without excessive compression, producing more stable behavior across evaluation metrics.

The choice of color representation further affects prediction performances from another perspective: although both hue and the A* channel encode the green-to-red color transition during growth, hue represents this change less distinctly and remains more sensitive to illumination effects. Consequently, hue-based signals deviate more from a smooth sigmoid growth trajectory, limiting the ability of prediction models to recover representative growth patterns and resulting in higher fitting and prediction errors.

Given the limited dataset size and variability in observation lengths (Wen, Verschoor, and Feng 2024), we primarily adopt the k-NN method and autoregressive prediction models, which naturally accommodate variable-length sequences. For comparison, we also trained a three-layer Temporal Convolutional Network (TCN) as a forecasting baseline. However, it performed poorly, likely because handling short sequences requires extensive input masking, which makes the learning more difficult. These results reinforce the effectiveness of more flexible or parametric approaches for better modeling variable-length temporal sequences for problems of this type.

Impact of scoring and objective functions

Tables 5 and 6 present how harvest scoring and objective-function choices affect optimization outcomes. To isolate the optimization logic, all plans are generated from sigmoid curves fitted to ground-truth growth observations, thereby removing prediction errors. We compared fuzzy and binary scoring, including variants with non-ripe and rot penalties, as showcased in the third subplot of Figure 4. For a fair comparison, all plans are evaluated using the same scoring method as in Table 3, which accounts for oversupply gains and penalties for harvesting too early or too late. In addition to the "overall score" in the tables, these two tables also report demand fulfillment, oversupply rate, and the number of fruits harvested

Table 4. Impact of color representation and normalization on growth prediction performances. Different color encoding and preprocessing strategies are compared to assess their impact on fitting and prediction. Evaluation includes the L2 distance for sigmoid curve fitting (when applicable) and prediction metrics consistent with Tables 1 and 2.

Data Source for Ripeness Evaluation	Method	L2 (fitting)	L2(pred.)	DTW(pred.)
A* value (Gray World)	s.k-NN☆	8.5	0.21	2.29
	Finetune GPT-2✪	–	0.17	2.30
	3-layer TCN	–	0.35	6.10
A* value (raw images)	s.k-NN	10.8	0.24	3.22
A* value (Intrinsic Image Decomp.)	s.k-NN	1.5	0.19	2.66
Hue value (Gray World)	s.k-NN	24.8	0.27	3.87

Table 5. Impact of scoring and objective-function choices on harvest-optimization performance under a daily demand of four fruits. Different combinations of options to oversupply, postpone harvest, and non-ripe and rot penalties are evaluated. The current harvest practice in the dataset and a simple "harvest once mature" policy are included as baselines, with their entries shown in *italic* in the first column. Performances are reported in terms of overall optimized score and demand fulfillment, consistent with Table 3. The table additionally reports the relative level of over-supply and the number of fruits whose harvests are scheduled too early or too late.

Scoring Function	Non-Ripe Penalty	Rot Penalty	D'_t	D_∞	Overall Score	Demand Fulfill.[%]	Over Supp.[%]	Late Harv.	Early Harv.
Current "Once Mature"	–	–	–	–	5.60	28.57	57.14	4	9
	–	–	–	–	8.22	50.00	71.43	0	0
Fuzzy	✓	✓	✓	✓	21.10	75.00	3.57	0	0
Fuzzy	✓	✓		✓	20.69	78.57	–	0	0
Fuzzy	✓	✓	✓		19.31	89.29	25.00	9	0
Fuzzy	✓	✓			19.54	100.00	–	6	0
Fuzzy	✓			✓	21.13	75.00	7.14	0	0
Fuzzy	✓			✓	20.03	75.00	–	0	0
Fuzzy	✓		✓		19.35	82.14	32.14	9	0
Fuzzy			✓	✓	21.02	78.57	3.57	1	0
Fuzzy					19.41	100.00	–	6	0
Binary			✓	✓	18.71	67.86	–	0	0
Binary			✓		16.11	92.86	21.43	13	0
Binary					17.15	100.00	–	9	0

Table 6. Impact of scoring and objective-function choices on harvest-optimization performance under a daily demand of two fruits. The same baseline, components, and evaluation metrics as in Table 5 are presented. Note that under this low-demand setting, harvest plans are largely unchanged by the dummy demand, since the limited weekly capacity naturally defers surplus fruits; such cases are indicated by bracketed checkmarks.

Scoring Function	Non-Ripe Panelty	Rot Penalty	D'_t	D_∞	Overall Score	Demand Fulfill.[%]	Over Supp.[%]	Late Harv.	Early Harv.
Current "Once Mature"	–	–	–	–	5.60	25.00	17.86	4	9
	–	–	–	–	8.22	46.43	17.86	0	0
Fuzzy	✓	✓	✓	✓	19.36	100.00	42.86	0	0
Fuzzy	✓	✓	✓		19.47	100.00	50.00	0	0
Fuzzy	✓	✓		✓	14.00	100.00	–	0	2
Fuzzy	✓	✓			14.00	100.00	–	0	2
Fuzzy	✓		✓	✓	18.83	100.00	35.71	0	0
Fuzzy	✓			(✓)	13.94	100.00	–	0	3
Fuzzy			✓	✓	19.36	100.00	42.86	0	0
Fuzzy				(✓)	13.94	100.00	–	0	3
Binary			✓	✓	18.83	100.00	35.71	0	0
Binary			✓		17.96	100.00	50.00	2	1
Binary				(✓)	13.83	100.00	–	0	3

too early or too late. Experiments in this subsection were conducted under two daily demand levels ($d = 2$ and $d = 4$), because although 32 fruits are observed, the actual harvest record shows that only 10 fruits are harvestable during that week.

As shown in Table 5, when the optimization is applied, including a non-ripe penalty is essential for controlling early harvest, particularly under high-demand conditions. Without this penalty, the optimizer exhibits "greedy" behavior, prioritizing immediate demand fulfillment whenever the reward is positive. This control is especially important when the oversupply allowance (D'_t) is enabled, as it prevents premature harvesting solely driven by positive scores.

Conversely, the rot penalty is essential to prevent excessive postponement, as illustrated in Table 6. Without this penalty and with the postpone option enabled, the algorithm tends to delay harvesting so that fruits approach their theoretical growth peaks, resulting in a relatively lower demand fulfillment rate. However, since natural growth eventually leads to decay, the rot penalty ensures the fruit is picked before quality degrades.

We also compared the effect of using binary scoring, which reflects the essence of the "harvest-once-mature" strategy while still operating within an optimization framework. Generally speaking, regardless of whether over-supply, postponed harvest, or penalty terms are applied, binary scoring consistently yields lower scores. The performance gap arises from the loss of temporal resolution when continuous changes during the growing period are reduced to binary scores, which limits the optimizer's ability to schedule harvest timing precisely.

In addition to comparing the schemes optimized with varying scoring functions, we also evaluated the optimization algorithms against the same baselines presented in Table 3: current real-world practice and a simple and non-optimized "harvest-once-mature" strategy. While the harvest-once-mature rule eliminates early and late harvesting relative to current practice, it generates substantial oversupply because it does not account for sequential demand across days. In contrast, the optimized approaches better coordinate harvest decisions over time, demonstrating the importance of handling plan optimization as an integrated temporal sequential problem rather than as independent daily targets.

Conclusion

This paper presents a novel end-to-end AI workflow, CROPI, for optimizing fruit-level harvest decisions. By quantifying and forecasting individual strawberry development, CROPI transforms raw in-field data into actionable decisions through four key components: objective ripeness (re-)definition, growth modeling, biological development forecasting, and daily harvest planning. Our experimental results show that integrating a color constancy algorithm, parametric growth function modeling, and assignment optimization is particularly effective for this type of data, providing high predictive accuracy while remaining computationally efficient. Together, these components form a data-driven bridge between biological observation and operational planning, shifting greenhouse management from reactive execution to proactive, optimized management in precision agriculture.

As a first step, we propose an objective, vision-based redefinition of fruit development that enables consistent growth and ripeness scoring. While image illumination can affect the accuracy of growth record extraction and prediction, this is an inherent challenge in real-world conditions. Hence, the most effective approach is to minimize such noise efficiently. In the strawberry case study, the A* channel of the CIELAB color space provides an intuitive and stable measure of ripening, while the Gray World algorithm normalizes illumination efficiently without introducing additional distortions. This combination directly captures the green-to-red transition and mitigates residual prediction errors. Beyond its use within CROPI, this ripeness definition can function as a standalone tool for more consistent quality assessment using vision systems or mobile devices.

The growth modeling and prediction in CROPI rely on continuous quality assessments of each fruit over its lifecycle. To isolate the evaluation of our models and methods, we utilized existing MOT annotations from Wen et al. (Wen, Verschoor, and Feng 2024), effectively assuming near-ideal perception conditions, and consequently reporting optimistic performances. While real-world tracking and perception noise inevitably introduce errors, the monotonic ripening process of strawberries – as they would only become redder – allows many inconsistencies to be detected and corrected during data preparation. Nevertheless, improving tracking accuracy would further benefit growth prediction and related precision agriculture tasks. Future work could explore how variations in tracking quality affect the practical deployment of growth models.

Furthermore, in CROPI, predicted fruit development is scored and incorporated as rewards in the optimization objective to guide harvest decisions. This formulation, however, is inherently crop-specific, as maturity standards and market preferences vary across species. In some cases, additional quality attributes such as size, firmness, sugar content, or post-harvest development must also be considered in forecasting and scoring (Jol, Wen, and Van Gemert 2023; Papageorgiou et al. 2018). Meanwhile, for climacteric fruits in particular, quality measures must account for post-harvest physiological changes. Consequently, extending the framework to other crops requires tailored scoring criteria, corresponding data analytics, and domain-specific knowledge supported by appropriate sensing and labeling strategies.

In the optimization phase, we employed fuzzy-logic scoring to evaluate rewards and optimize harvest plans. By preserving the temporal resolution of fruit development, fuzzy scoring allows the optimizer to schedule harvests with a precision that binary "harvest-once-mature" rules cannot capture. We incorporated explicit penalties to regulate harvesting timing, controlling the over-use of early harvesting driven by short-term rewards and excessive postponement that conflicts with natural quality decay. While current parameters were self-defined to establish a demonstration, transitioning this framework into a production-ready tool necessitates a human-in-the-loop calibration. This allows growers and facility managers to align the thresholds and weights to reflect cultivar traits, operational constraints, and market priorities. Moreover, incorporating stochastic or robust optimization into the scoring or optimization stage represents a potential direction for future work, particularly for accommodating risk-averse decision-making in harvest planning.

Our demonstration included benchmarks on three public datasets, including two for ripeness grading (Elhariri, El-Bendary, and Mahmoud Saleh 2023; Wen et al. 2023) (see Processing the images) and one for growth monitoring (Wen, Verschoor, and Feng 2024) (see Modeling the growths). Due to limited open-source data, we relied on nearby climate data (Hersbach et al. 2023) to approximate on-site microclimate conditions, and we used a virtual price (in the reward function) and demand profiles in the planning stage. These proxies are sufficient to illustrate the workflow functionality, particularly since price and demand are generally stable under pre-defined contracts. However, full-scale deployment would require site-specific data and realistic market information to fine-tune model parameters and objective functions. Our validation thus demonstrates proof-of-concept functionality rather than operational readiness. While the workflow is designed to be broadly applicable, accurate and reliable deployment will require extensive site-specific testing and calibration under varying environmental conditions. Furthermore, the lack of comprehensive open-source datasets is a shared concern for benchmarking how AI can be applied in agriculture (Maya Gopal and Bhargavi 2019; Wen, Verschoor, and Feng 2024). We hope the introduction of this novel approach also motivates the curation of more comprehensive datasets to support advanced learning and optimization paradigms, such as deep reinforcement learning and predict-then-optimize strategies.

In summary, CROPI demonstrates the concept of transforming diverse temporal information into a unified form suitable for optimization, by mining and harmonizing meaningful patterns from the diverse data sources. With the rapid assignment optimization enabled by the Hungarian method, CROPI supports real-time planning and advances intelligent greenhouse management and more efficient harvesting operations. This capability could be further enhanced with adaptive control or reinforcement learning, where transformer-based growth models are better suited due to their flexible architecture and ability to capture complex temporal dependencies, enabling more accurate short- and medium-term predictions. At the same time, traditional parametric models, such as k-NN, provide competitive performance with simpler machine learning methods, preserving optimization quality while significantly reducing computational cost, making them ideal for routine, real-time deployment.

Ultimately, the core strength of this framework lies in its pragmatic approach to precision agriculture. Because the entire CROPI workflow relies exclusively on nondestructive, in-field data, it avoids the bottleneck of destructive sampling. This makes the system foundational and pragmatic for precision agriculture and autonomous farming, offering a scalable approach toward data-driven, optimized harvest management.

Acknowledgements

We gratefully acknowledge the generous support and funding of this project. We also thank our colleagues and collaborators involved in the wider project for their support and contributions. Finally, we are grateful to Kim for proofreading and providing helpful suggestions.

Author contributions

CRediT: **Junhan Wen:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **Thomas Abeel:** Funding acquisition, Supervision, Writing – review & editing; **Mathijs de Weerdt:** Funding acquisition, Project administration, Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project is fully funded by Topsector Tuinbouw & Uitgangsmaterialen (The Netherlands), Innovatiefonds Hagelunie, and Interpolis.

ORCID

Junhan Wen http://orcid.org/0009-0007-4295-2682
Thomas Abeel http://orcid.org/0000-0002-7205-7431
Mathijs de Weerdt http://orcid.org/0000-0002-0470-6241

Data availability statement

The data that support the findings of this study are: (i) "Strawberry-DS": openly available in *Mendeley Data* at http://doi.org/10.17632/z6dtfdpzz8.1, reference number (Elhariri, El-Bendary, and Mahmoud Saleh); (ii) "Data underlying the research of strawberry quality prediction with infield data": openly available in *4TU.ResearchData* at http://doi.org/10.4121/96a1cb2d-9f16-470f-8c46-744d5985a140.v1, reference number (Wen et al. 2023); (iii) "The Growing Strawberries" Dataset: openly available in *4TU.ResearchData* at http://doi.org/10.4121/e3b31ece-cc88-4638-be10-8ccdd4c5f2f7.v2, reference number (Wen, Verschoor, and Feng 2024); (iv) "ERA5 hourly data on single levels from 1940 to present": openly available in *Climate Data Store* at http://doi.org/10.24381/cds.adbb2d47, reference number (Hersbach et al. 2023).

References

Achiam, Josh, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, and Shyamal Anadkat. 2023. "GPT-4 Technical Report." *arXiv preprint*.

Ahmed, Shifaz, Pelletier Charlotte, François Petitjean, and I Webb. Geoffrey. 2023. "Elastic Similarity and Distance Measures for Multivariate Time Series." *Knowledge and Information Systems* 65 (6): 2665–2698.

Ayala-Silva, Tomas, Raymond J Winterstein, Alan W Meerow, Michael Winterstein, Cuauhtemoc Cervantes, and J. Steven Brown. 2005. "Determination of Color and Fruit Traits of Half-Sib Families of Mango (Mangifera Indica l.)." *Proceedings of the Florida State Horticultural Society* 118: 253–257.

Bali, Nishu, and Anshu Singla. 2021. "Deep Learning Based Wheat Crop Yield Prediction Model in Punjab Region of North India." *Applied Artificial Intelligence* 35 (15): 1304–1328. https://doi.org/10.1080/08839514.2021.1976091.

Bellemare, Marc F., and Jeffrey R. Bloem. 2018. "Does Contract Farming Improve Welfare? A Review." *World Development* 112:259–271. ISSN 0305-750X https://doi.org/10.1016/j.worlddev.2018.08.018.

Cacchiani, Valentina, Alberto Caprara, and Paolo Toth. 2010. "Solving a Real-World Train-Unit Assignment Problem." *Mathematical Programming* 124 (1–2): 207–231. https://doi.org/10.1007/s10107-010-0361-y.

Cai, Wanlin, Yuxuan Liang, Xianggen Liu, Jianshuai Feng, and Yuankai Wu. 2024. "Msgnet: Learning Multi-Scale Inter-Series Correlations for Multivariate Time Series Forecasting." *Proceedings of the AAAI Conference on Artificial Intelligence* 38: 11141–11149.

Catalá, Luis P, M. Susana Moreno, Aníbal M Blanco, and J. Alberto Bandoni. 2016. "A Bi-Objective Optimization Model for Tactical Planning in the Pome Fruit Industry Supply Chain." *Computers and Electronics in Agriculture* 130:128–141. https://doi.org/10.1016/j.compag.2016.10.008.

CTIFL. "Code couleur fraise." Accessed May 2025. https://www.ctifl.fr/code-couleur-fraise.

Cuturi, Marco, and Mathieu Blondel. 2017. "Soft-Dtw: A Differentiable Loss Function for Time-Series." *International conference on machine learning*, Sydney, Australia, 894–903. PMLR.

Das, Jnaneshwar, Gareth Cross, Chao Qu, Anurag Makineni, Pratap Tokekar, Yash Mulgaonkar, and Vijay Kumar. 2015. "Devices, Systems, and Methods for Automated Monitoring Enabling Precision Agriculture." In 2015 IEEE international conference on automation science and engineering (CASE), Gothenburg, Sweden, 462–469. IEEE.

Elhariri, Esraa, Nashwa El-Bendary, and Samir Mahmoud Saleh. 2023. "Strawberry-Ds: Dataset of Annotated Strawberry Fruits Images with Various Developmental Stages." *Data in Brief* 48: 109165. https://doi.org/10.1016/j.dib.2023.109165.

Fang, Shih-Lun, Yu-Hsien Kuo, Le Kang, Chu-Chung Chen, Chih-Yu Hsieh, Min-Hwi Yao, and Bo-Jein Kuo. 2022. "Using Sigmoid Growth Models to Simulate Greenhouse Tomato Growth and Development." *Horticulturae* 8 (11): 1021. https://doi.org/10.3390/horticulturae8111021.

Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Gerhard, Daniel, and Elena Moltchanova. 2022. "A Richards Growth Model to Predict Fruit Weight." *Australian & New Zealand Journal of Statistics* 64 (4): 413–421. https://doi.org/10.1111/anzs.12380.

Gijsenij, Arjan, and Theo Gevers. 2011. "Color Constancy Using Natural Image Statistics and Scene Semantics." *IEEE Transactions on Pattern Analysis & Machine Intelligence* 33 (4): 687–698. https://doi.org/10.1109/TPAMI.2010.93.

Goatman, Keith A, A. David Whitwam, A. Manivannan, John A Olson, and Peter F Sharp. 2003. "Colour Normalisation of Retinal Images." In *Proceedings of medical image understanding and analysis*, Sheffield, United Kingdom, 49–52. The University of Sheffield United Kingdom.

Gómez-Lagos, Javier E, Marcela C González-Araya, Wladimir E Soto-Silva, and Masly M Rivera-Moraga. 2021. "Optimizing Tactical Harvest Planning for Multiple Fruit Orchards Using a Metaheuristic Modeling Approach." *European Journal of Operational Research* 290 (1): 297–312. https://doi.org/10.1016/j.ejor.2020.08.015.

Harel, Ben, Yael Edan, and Yael Perlman. 2022. "Optimization Model for Selective Harvest Planning Performed by Humans and Robots." *Applied Sciences* 12 (5): 2507. https://doi.org/10.3390/app12052507.

Hassanjani Roshan, Shiva, Javad Kazemitabar, and Ghorban Kheradmandian. 2022. "Artificial Intelligence Aided Agricultural Sensors for Plant Frostbite Protection." *Applied Artificial Intelligence* 36 (1): 2031814. https://doi.org/10.1080/08839514.2022.2031814.

Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas. 2023. "ERA5 Hourly Data on Single Levels from 1940 to Present." Accessed 15 July, 2024.

Hopf, Alwin, Kenneth J. Boote, Juhyun Oh, Zhengfei Guan, Shinsuke Agehara, Vakhtang Shelia, Vance M. Whitaker, Senthold Asseng, Xin Zhao, and Gerrit Hoogenboom. 2022. "Development and Improvement of the Cropgro-Strawberry Model." *Scientia Horticulturae* 291:110538. ISSN 0304-4238. https://doi.org/10.1016/j.scienta.2021.110538.

Hueth, Brent, Ethan Ligon, Steven Wolf, and Steven Wu. 1999. "Incentive Instruments in Fruit and Vegetable Contracts: Input Control, Monitoring, Measuring, and Price Risk." *Applied Economic Perspectives and Policy* 21 (2): 374–389. https://doi.org/10.2307/1349886.

Jannatizadeh, Abbasali, Mehdi Rezaei, Abbas Rohani, Shaneka Lawson, and Reza Fatahi. 2023. "Towards Modeling Growth of Apricot Fruit: Finding a Proper Growth Model." *Horticulture, Environment, and Biotechnology* 64 (2): 209–222. https://doi.org/10.1007/s13580-022-00475-x.

Jha, S. N., Sangeeta Chopra, and A. R. P. Kingsly. 2007. "Modeling of Color Values for Nondestructive Evaluation of Maturity of Mango." *Journal of Food Engineering* 78 (1): 22–26. https://doi.org/10.1016/j.jfoodeng.2005.08.048.

Jol, Cees, Junhan Wen, and Jan Van Gemert. 2023. "Non-Destructive Infield Quality Estimation of Strawberries Using Deep Architectures." In *Proceedings of the IEEE/CVF International Conference on Computer Vision* Paris, France, 515–524.

Karpathy, Andrej 2022. NanoGPT. https://github.com/karpathy/nanoGPT.

Keogh, E., and C. A. Ratanamahatana. 2005. "Exact Indexing of Dynamic Time Warping." *Knowledge and Information Systems* 7 (3): 358–386. https://doi.org/10.1007/s10115-004-0154-9.

Kuhn, Harold W 1955. "The Hungarian Method for the Assignment Problem." *Naval Research Logistics Quarterly* 2 (1–2): 83–97. https://doi.org/10.1002/nav.3800020109.

Lavanya, M., and B. Santhi. 2021. "Hungarian Optimization Technique Based Efficient Resource Allocation Using Clustering Unbalanced Estimated Cost Matrix." *Journal of Ambient Intelligence and Humanized Computing* 12 (5): 5525–5540. https://doi.org/10.1007/s12652-020-02063-2.

Lee, Dah-Jye, James K. Archibald, and Guangming Xiong. 2011. "Rapid Color Grading for Fruit Quality Evaluation Using Direct Color Mapping." *IEEE Transactions on Automation Science and Engineering* 8 (2): 292–302. https://doi.org/10.1109/TASE.2010.2087325.

Li, Tingpeng, Yue Li, and Yanling Qian. 2016. "Improved Hungarian Algorithm for Assignment Problems of Serial-Parallel Systems." *Journal of Systems Engineering and Electronics* 27 (4): 858–870.

Li, Zhengqi, and Noah Snavely. 2018. "Learning Intrinsic Image Decomposition from Watching the World." In *Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 9039–9048. https://doi.org/10.1109/CVPR.2018.00942.

Mavridou, Efthimia, Eleni Vrochidou, George A Papakostas, Theodore Pachidis, and Vassilis G Kaburlasos. 2019. "Machine Vision Systems in Precision Agriculture for Crop Farming." *Journal of Imaging* 5 (12): 89. https://doi.org/10.3390/jimaging5120089.

Maya Gopal, P. S., and R. Bhargavi. 2019. "Performance Evaluation of Best Feature Subsets for Crop Yield Prediction Using Machine Learning Algorithms." *Applied Artificial Intelligence* 33 (7): 621–642.

Ni Made Defy Janurianti, I Made Supartha Utama, and Ida Bagus Wayan Gunam. 2021. "Colour and Quality of Strawberry fruit (Fragaria x ananassa Duch.) at Different Levels of Maturity." *SEAS (Sustainable Environment Agricultural Science)* 5 (1): 22–28. https://doi.org/10.22225/seas.5.1.3166.22-28.

Omar, Ahumada, and J. Rene Villalobos. 2009. "Application of Planning Models in the Agri-Food Supply Chain: A Review." *European Journal of Operational Research* 196 (1): 1–20. ISSN 0377-2217. https://www.sciencedirect.com/science/article/pii/S0377221708001987.

Otsuka, Keijiro, Yuko Nakano, and Kazushi Takahashi. 2016. "Contract Farming in Developed and Developing Countries." *Annual Review of Resource Economics* 8 (1): 353–376. https://doi.org/10.1146/annurev-resource-100815-095459.

Papageorgiou, E. I., K. Aggelopoulou, T. A. Gemtos, and G. D. Nanos. 2018. "Development and Evaluation of a Fuzzy Inference System and a Neuro-Fuzzy Inference System for Grading Apple Quality." *Applied Artificial Intelligence* 32 (3): 253–280. https://doi.org/10.1080/08839514.2018.1448072.

Pattnaik, Gayatri, Vimal K. Shrivastava, and K. Parvathi. 2020. "Transfer Learning-Based Framework for Classification of Pest in Tomato Plants." *Applied Artificial Intelligence* 34 (13): 981–993. https://doi.org/10.1080/08839514.2020.1792034.

Qiu Nate Gruver, Shikai, Marc Finzi, and Andrew Gordon Wilson. 2023. "Large Language Models are Zero Shot Time Series Forecasters." In *Advances in Neural Information Processing Systems 36*. New Orleans, Louisiana, USA.

Richards, Francis J 1959. "A Flexible Growth Function for Empirical Use." *Journal of Experimental Botany* 10 (2): 290–301. https://doi.org/10.1093/jxb/10.2.290.

Sakoe, Hiroaki, and Seibi Chiba. 1978. "Dynamic Programming Algorithm Optimization for Spoken Word Recognition." *IEEE Transactions on Acoustics, Speech & Signal Processing* 26 (1): 43–49. https://doi.org/10.1109/TASSP.1978.1163055.

Sala, Florin, Cosmin Alin Popescu, Mihai Valentin Herbei, and Ciprian Rujescu. 2020. "Model of Color Parameters Variation and Correction in Relation to 'Time-View' Image Acquisition Effects in Wheat Crop." *Sustainability* 12 (6): 2470. https://doi.org/10.3390/su12062470.

Schanda, Janos 1996. "CIE Colorimetry and Colour Displays." In *Color and Imaging Conference*, November 1996, Scottsdale, Arizona, USA, 230–234.

Taşkıner, Tuğçe, and Bilge Bilgen. 2021. "Optimization Models for Harvest and Production Planning in Agri-Food Supply Chain: A Systematic Review." *Logistics* 5 (3): 52. ISSN 2305-6290. https://doi.org/10.3390/logistics5030052.

Wen, Junhan, Mathijs de Weerdt, Thomas Abeel, Camiel Verschoor, Lisanne Schuddebeurs, Klaas Walraven, Stijn Jochems, and Vera Theelen. 2023. "Data Underlying the Research of Quality Prediction of Strawberries with Rgb Image Segments." https://doi.org/10.4121/96a1cb2d-9f16-470f-8c46-744d5985a140.

Wen, Junhan, Camiel R. Verschoor, Chengming Feng, Epure Irina-Mona, Abeel, Thomas, de Weerdt, Mathijs. 2024. "The Growing Strawberries Dataset: Tracking Multiple Objects with Biological Development Over an Extended Period." IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, Hawaii, USA, 7089–7099. https://doi.org/10.4121/e3b31ece-cc88-4638-be10-8ccdd4c5f2f7.v2.

Zhang, Xiyuan, Ranak Roy Chowdhury, Rajesh K Gupta, and Jingbo Shang. 2024. "Large Language Models for Time Series: A Survey." *arXiv preprint* 2402.01801.