Proposal for a System to Estimate the Best Time to See Yellow Leaves Using IoT Devices for Tourists

Takumi Kitayama
Electronic Information Course
Polytechnic University
Kodaira-shi, Tokyo, Japan
b21306@uitec.ac.jp

Kenji Terada, Masaki Endo,
Tsuyoshi Tanaka,Shigeyoshi Ohno
Division of Core Manufacturing
Polytechnic University
Kodaira-shi, Tokyo, Japan
e-mail: {k-terada, endou, t-tanaka, ohno}@uitec.ac.jp

Hiroshi Ishikawa
Graduate School of Systems Design
Tokyo Metropolitan University
Hino-shi, Tokyo, Japan
e-mail: ishikawa-hiroshi@tmu.ac.jp

***Abstract*—Japan's diverse topography results in distinct seasonal landscapes, which are a significant draw for tourists. Among the numerous natural attractions, the autumnal coloration of foliage is popular. However, the phenological timing of leaf senescence and abscission varies considerably across different geographical locations. Consequently, visitors should be disappointed by arriving either before the onset of yellowing or after leaf fall. To address this issue, we propose a system for the estimation of autumnal leaf coloration timing. The system leverages Internet of Things (IoT) devices to collect environmental data, including photos of ginkgo trees and insolation conditions. This data is subsequently processed to predict the near-future timing of leaf yellowing and subsequent abscission. A prototype demonstration system was developed, and its effectiveness in predicting the target phenological events was confirmed through empirical evaluation.**

Keywords- Yellow Leaves Tourism; Internet of Things; Artificial Intelligence; Estimating.

I. INTRODUCTION

Biological phenological observations [1] rely on visual observation, so the huge human cost is an issue. It has also been pointed out that in urban environments, the shading environment caused by buildings may affect biological phenology. However, there are few studies on biological phenological observations [2][3][4] of local plants in urban environments that reduce human costs.

To solve these issues, we propose a method to reduce human costs and to determine and predict the yellowing of local and specific individual ginkgo trees in urban environments. First, we used an IoT(Internet of Things) device we constructed to automatically collect fixed-point photographs of ginkgo trees and meteorological data around the ginkgo trees. Furthermore, the number of yellow leaf pixels is extracted from the ginkgo image to calculate the degree of yellowing. The color of ginkgo leaves in an image changes depending on factors such as cloud cover, camera performance, and the amount of leaves. Simply extracting the number of pixels cannot handle these changes, and areas that appear blue or green may be judged to be yellow leaf pixels. For this reason, we decided to use deep learning to identify colors in response to changes in the image. Furthermore, focusing on the fact that natural phenomena, such as leaves turning yellow or falling, are not binary values of whether or not the leaves have turned yellow or fallen, we aim to predict the ginkgo phenology in more detail by extracting indicators that show the degree of yellowing and the degree of leaf fall, and performing regression analysis on each as the objective variable.

This paper is organized as follows: Section 2 introduces related research. Section 3 explains the observation system built using IoT devices. Section 4 explains the method for analyzing the collected observation data. Section 5 presents and discusses the results of the analysis. Section 6 concludes this paper and describes future issues.

II. RELATED RESEARCH

As biological phenological observations have been reduced, various studies on biological phenological observations have been reported to solve the problem. Below, we will discuss research related to the development of biological phenological observation methods.

In Endo et al.'s research [5][6], we proposed a method to estimate the timing of relic season changes in biological phenological observations at low cost from X (formerly Twitter) location-attached posts. By analyzing the names of organisms such as ginkgo and maple in the posts and co-occurring words indicating their location and state, the timing of biological phenological changes was estimated from the frequency of posts. Furthermore, the effectiveness of the proposed method was verified by comparing with observation data from the Japan Meteorological Agency.

In Iha et al.'s research [7], we used post data related to cherry blossoms from March to the end of April 2022 as a dataset and performed time series prediction of the number of posts using machine learning. As a result, we confirmed an improvement in the precision and recall of the time series prediction model of the number of posts compared to conventional methods.

In Ito et al.'s research [8], they developed a robot that can automatically measure plant growth information by utilizing low-cost IoT devices and open source image processing libraries. This robot was used to periodically capture images of spinach growth, demonstrating its potential for application in growth prediction and detection of poor growth.

In a study by Sato et al. [9], multispectral observations using a drone and IoT devices were used to observe the growth status of wheat using vegetation indices.

As described above, many methods have been researched for efficiently observing biological phenologies and plants using SNS(Social Networking Service) and IoT devices, but there has been no research on a system that can estimate the best time for yellow leaves to appear.

III. Observation Data

Biological phenological observations were performed in accordance with the Japan Meteorological Agency's biological phenological observation guidelines, and fixed-point photography was performed on ginkgo trees at the Polytechnic University as specimen trees. In addition, meteorological information from the surrounding area (hereinafter referred to as sensor measurements) is measured as a feature used to predict the yellowing and falling of ginkgo leaves. The sensor measurements are temperature, humidity, air pressure, carbon dioxide concentration, and illuminance.

The observation period is from November 1, 2024 to January 10, 2025. The measurement frequency was one image and sensor measurement value set per minute. However, each sensor measurement value was taken for 24 hours, but images were taken only from 6:00 to 18:00.

IV. Observation System

The measurement system to be constructed is shown in Figure 1. Ginkgo measurements will be made using products from M5Stack. Fixed-point photographs of the ginkgo trees will be taken using a Timer Camera. Temperature, humidity, and air pressure will be measured using a HAT-YUN, CO2 using an SGP30, and illuminance using a BH1750FVI-TR. The sensor measurements will be collected on the M5Stick microcontroller.

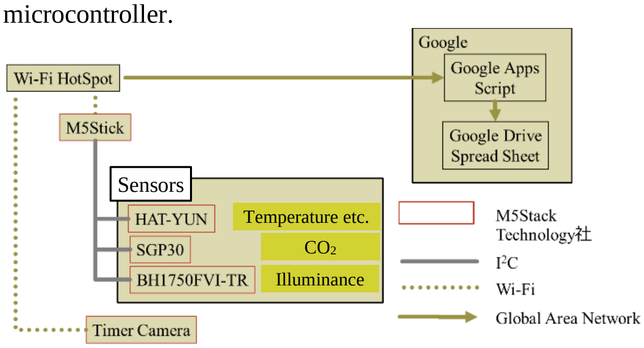

Figure 1. Overview of the observation system and data flow.

Images of the ginkgo trees will be saved to Google Drive and the sensor measurements will be saved to a spreadsheet using a script written in Google Apps Script provided by Google.

V. Analysis Method

From the acquired image, an area that only contains ginkgo leaves (hereafter, ginkgo image) is cut out. Furthermore, the ginkgo image is divided into 10x10 pixel images (hereafter, square images), and each square image is classified into "green", "yellow", and "other". The classification method is to first select only ginkgo images at times when the illuminance, one of the sensor measurement values, is between 1000 and 10000lx. These images are divided into square images and labeled as "green", "yellow", and "other". These square images are used as learning data for training, and a model is generated that classifies the images into three classes: "green", "yellow", and "other". When an image containing only ginkgo leaves is divided into 10x10 pixels using this model, the number of images classified into each class is counted. In calculating the index, the ratio of the number of yellow class classifications to the total number of green and yellow class classifications (hereafter referred to as leaf amount), which indicates the entire ginkgo leaf, is defined as the "yellow leaf rate," and the "defoliation rate" is defined as the rate at which the leaf amount at the time of measurement has decreased from the maximum leaf amount obtained up to the time of measurement (hereafter referred to as maximum leaf amount).

According to the observation conditions of ginkgo in the biological phenology observation of the Japan Meteorological Agency, the yellow leaf day refers to the first day when the majority of the leaves have turned yellow when viewed as a whole and almost no green parts are visible. The defoliation day refers to the first day when approximately 80% of the leaves of the specimen tree have fallen. Therefore, the leaves are judged to be yellow when the yellow leaf rate is 80% or more of the maximum leaf amount, and the leaves are judged to be fallen when the defoliation rate is 80% or more.

Multivariate time series prediction is performed using LightGBM. Accuracy is verified using the yellow leaf rate and defoliation rate as the objective variables, and sensor measurements and processed data from them as explanatory variables.

VI. Results and Discussion

The results of three-class image classification using the image classification model ResNeXt are shown below. The three classes are yellow, which means yellow leaves, green before the leaves turn yellow, and other colors, which mean fallen leaves. Evaluation data was classified using the model determined to be the best by generalized k (k=5)-fold cross-validation, and the evaluation results shown in Table 1 were obtained. According to Table 1, the precision rate for the green class classification is a little low at 0.892. However, all other colors were above 0.92. We counted the number of areas that were green, yellow, and other colors in the ginkgo image, and we believe that we were able to calculate the indices of yellow leaf rate and fallen leaf rate with high accuracy.

TABLE I. Evaluation of Image Classification of Square Images

Class Name	Precision	Recall	F-Value
Green	0.892	0.928	0.909
Yellow	0.945	0.927	0.935
Others	0.921	0.928	0.924

Details of the precision rate are shown in Table 2. The precision rate is an evaluation index that indicates how accurate the prediction was. Looking at Table 2, we see that the proportion of images classified as green that were actually labeled as yellow was 0.082, and the proportion of images labeled as other was 0.027. In other words, there were more images erroneously predicted to be green that were labeled as yellow than as other. This suggests that while the system was relatively accurate in classifying images where the correct answer was other classes such as trunks and branches, it is possible that the classification of green and yellow classes did not capture the subtle changes that occur when leaves change from green to yellow. In particular, when shadows were cast on the specimen tree due to the way the sunlight hit it after the day the leaves turned yellow, there were many images in which those parts were erroneously classified as green. Therefore, we believe that accuracy can be improved by devising a shooting method and image processing method that are not affected by the direction of the sun or shadows.

TABLE II. DETAILS OF PRECISION RATE

		Classification results		
		Green	Yellow	Others
Actual	Green	0.892	0.928	0.909
	Yellow	0.945	0.927	0.935
	Others	0.921	0.928	0.924

TABLE III. OBJECTIVE AND EXPLANATORY VARIABLES OF THE PREDICTION MODEL

Objective variables	Explanatory variables
Yellowing and fallen leaf rates three days later	Average yellowing rate for the past three days
	Average falling leaf rate for the past three days
	Illuminance
	Average of illuminance and CO_2 integrated value for three days
	Additional value of average illuminance from 6:00 to 18:00 on the same day

The objective variables and explanatory variables used in the LightGBM analysis were defined as shown in Table 3 for the leaf yellowing rate prediction model and leaf fall rate prediction model. Note that temperature, humidity, and air pressure, which were planned to be used as explanatory variables, were not used as explanatory variables because only fixed values were recorded from the middle of the observation period. For the explanatory variables in Table 3, the objective variables were predicted based on the average leaf yellowing rate over the past three days and the average leaf fall rate over the past three days. In addition, the objective variables were predicted based on the average illuminance and the integrated value of illuminance and CO_2 over three days, as sunlight and photosynthetic activity would affect the objective variables. Furthermore, the average illuminance value from 6:00 to 18:00 on the same day was added to replace the integrated temperature to improve the prediction accuracy. Regarding the objective variables, the prediction period for yellowing and leaf fall judgment in previous studies was three days later, so the objective variables in this study were set to the leaf yellowing rate and leaf fall rate three days later from the last day of the average value of the past three days. Note that these data are saved every minute. Therefore, for the data at 12:00 on November 10, 2024, the average value of the explanatory variables over the past three days is the average value of the data from 11:59 on November 8 to 11:59 on November 10, 2024, and the target variables are the yellowing and falling leaf rates at 12:00 on November 13, 2024.

Figure 2 shows the actual and predicted values of the yellowing leaf rate, and Figure 3 shows the actual and predicted values of the defoliation rate. Additionally, Table 4 shows an evaluation of the yellowing leaf rate prediction model and the defoliation rate prediction model.

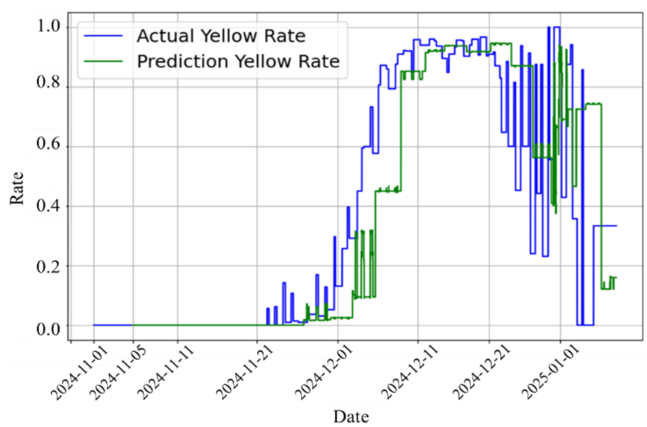

Figure 2. Actual value (blue) and predicted value (green) of the yellowing rate.

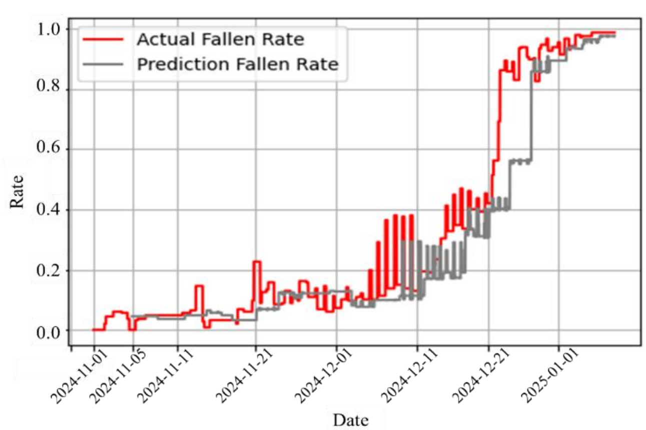

Figure 3. Actual value (red) and predicted value (gray) of leaf fall rate.

 ISBN: 978-1-68558-279-1

Figures 2 and 3 show that the predicted values of the yellowing and falling leaf rates fluctuate about five days later than those calculated from the ginkgo image. In addition, after the fluctuation, there is no fluctuation and the rate remains flat for about three days. The reason for the approximately five-day delay in the fluctuation is thought to be that the data learned when the model output the predicted value of the evaluation data was from five days ago, so it was not possible to predict it as time-series data. In addition, the reason for the leveling off is that the yellowing and falling leaf rates were restricted to extract only the ginkgo image and only the time when the illuminance was within a certain range, leaving the yellowing and falling leaf rates blank for the time when the illuminance was outside the certain range. To fill this gap, the leaf yellowing rate and leaf fall rate from the most recent time when the illuminance was within a certain range were used. As a result, while the explanatory variables fluctuated during the learning data period, the objective variables, the leaf yellowing rate and leaf fall rate, did not fluctuate and remained flat during the period when the illuminance was outside of a certain range, so it is thought that the predicted values also produced similar outputs.

TABLE IV. RMSE OF EACH REGRESSION MODEL

Model	RMSE
The yellowing leaf rate prediction model	0.164
The defoliation rate prediction model.	0.083

From Table 4, an RMSE(Root Mean Squared Error) of 0.164 was calculated for the leaf yellowing rate prediction model. RMSE is an index that shows the difference between the predicted value and the actual leaf yellowing rate or leaf fall rate. This means that there is an average error of 0.164 between the predicted value and the actual value. A leaf yellowing day requires a leaf yellowing rate of 80% or more, but even if a predicted value of 0.8 is obtained, there is an average error of ±0.164, so the actual value varies from 0.636 to 0.964. However, in Figures 2 and 3, the leaf yellowing rate fluctuates wildly from the time of the leaf fall day, and it is possible that the influence of this period affects the RMSE of the leaf yellowing rate prediction model.

VII. CONCLUSION AND FUTURE WORK

In this study, we first proposed a method to calculate the degree of yellowing and defoliation from ginkgo images using image classification by deep learning. While changes in the date and amount of leaves during the observation period affect the color of the leaves in ginkgo images, we were able to extract the yellowing and defoliation rates from ginkgo images by generating a highly accurate image classification model.

Next, we verified a method for predicting the yellowing and defoliation rates three days later using processed weather data as explanatory variables.

Future challenges for observation methods using IoT devices are as follows. First, we need to improve the accuracy of image classification by devising a shooting method that is not affected by the direction of the sun or shadows. Next, there will be the detection of anomalies in meteorological data measurements and the response to sensor anomalies through sensor multiplexing.

In this study, there were malfunctions with the temperature, humidity, and air pressure sensors, and their measurements could not be used. If these sensor values could be used in addition to illuminance and carbon dioxide concentration data, it is expected that prediction accuracy will improve. In particular, air pressure affects the weather, so it should augment the illuminance sensor value.

Future challenges in the method of acquiring the rate of yellowing and defoliation from ginkgo images are as follows. First, the dataset of the rate of yellowing and defoliation will be enriched by devising an image processing method that makes it possible to read color from ginkgo images that are either too bright or too dark.

In this study, images were only taken from 6:00 to 18:00, so a future challenge is how to obtain changes in image data during this period. In addition, the measurement location was fixed at one place, so in the future it will be necessary to verify whether this method is effective for trees in different locations.

By making improvements to these issues, we hope to aim to automatically predict when the leaves will yellow and defoliate for specimen ginkgo trees.

REFERENCES

[1] Japan Meteorological Agency. [Online]. Available from: https://www.jma.go.jp/jma/indexe.html [retrieved: 1, 2023].

[2] Y. Mizutani and K. Yamamoto, “A sightseeing spot recommendation system that takes into account the change in circumstances of users,” ISPRS International Journal of Geo-Information, vol. 6, no. 10, pp. 303, 2017.

[3] Y. Wang and W. Yue, “Proposal and Evaluation of a Pictorial Map Generation Method Based on the Familiarity and Satisfaction of Sightseeing Spots Calculated from Photos Posted on SNS (in Japanese),” Journal of the Japan Personal Computer Application Technology Society, vol. 16, no. 1, pp. 1–10, 2021.

[4] G. Fang, S. Kamei, and S. Fujita, “How to extract seasonal features of sightseeing spots from Twitter and Wikipedia (preliminary version),” Bulletin of Networking, Computing, Systems, and Software, vol. 4, no. 1, pp. 21–26, 2015.

[5] M. Endo, M. Takahashi, M. Hirota, M. Imamura, and H. Ishikawa, “Analytical Method using Geotagged Tweets Developed for Tourist Spot Extraction and Real-time Analysis,” International Journal of Informatics Society (IJIS), vol. 12, no. 3, pp. 157–165, 2021.

[6] M. Takahashi, M. Endo, S. Ohno, M. Hirota, and H. Ishikawa, “Automatic detection of tourist spots and best-time estimation using social network services,” International Workshop on Informatics 2020, pp. 65–72, 2020.

[7] A. Iha, K. Terada, M. Endo, T. Tanaka, S. Ohno, and H. Ishikawa, “Relation between theTtime-series Forecasting Methods and the Best Times to See Cherry Blossoms Estimation,” 17th International Workshop on Informatics (IWIN2023), pp.145-152, 2023.

[8] J. Ito, et al. "Development and Performance Evaluation of a Plant Phenotyping Platform Using Low-cost IoT Devices," Agricultural Information Research 30 (2), pp.13–23, 2021.

[9] T. M. Sato, J. Kurihara, T. Yumura, Y. Kakinami, "Predicting Optimal Harvest Time for Wheat Based on Multispectral Drone Images and Meteorological Data Collected Using IoT Sensors," Journal of Hokkaido Information University, vol.36, pp39-52, 2024.