International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 9s, 2025
https://www.theaspd.com/ijes.php

Automated Image Recognition of Staminate and Pistillate Flowers in Cucurbit Crops for Precision Pollination and Growth Management

Kuo-Dung Chiou[1], Shin-Hau Chiou[2], Yi-Zhen Chen[3], Yu-Shen Liang[4] and Chia-Ying Chang[5]*
[1] Fengshan Tropical Horticultural Experiment Branch, Taiwan Agricultural Research Institute, Ministry of Agriculture, Kaohsiung, Taiwan and Department of Plant Industry, National Pingtung University of Science and Technology, Pingtung, 91201, Taiwan.
[2,3,5]Bachelor of Program in Scientific Agriculture, National Pingtung University of Science and Technology, Pingtung, 91201, Taiwan
[4]Department of Plant Industry, National Pingtung University of Science and Technology, Pingtung 91201, Taiwan
[1]kuodung@tari.gov.twl, [2]abeth20050102@gmail.com, [3]zhen9210@gmail.com,
[4]justinliang@mail.npust.edu.tw and [5]csiefly@mail.npust.edu.tw

Abstract: *Cucurbit crops, such as watermelon, pumpkin, and cucumber, exhibit distinct monoecious characteristics, with significant differences in the timing and morphology of staminate and pistillate flowers. Staminate flowers are typically smaller, borne on shorter pedicels, and characterized by prominent anther structures, whereas pistillate flowers are larger, borne on longer pedicels, and distinguished by well-developed ovary structures. Image recognition technology enables automated flower identification and classification, assisting farmers in optimizing pollination strategies and crop management. In controlled environments, manual pollination remains a critical step for enhancing fruit quality and yield. Automated monitoring of flowering stages and floral abundance allows for precise control of production timing and output. Traditional manual pollination relies heavily on human experience, which is time-consuming, error-prone, and inefficient due to challenges in visual identification, time constraints, and precision requirements. By integrating an image-based recognition system, the identification of staminate and pistillate flowers can be automated, streamlining pollination processes, improving efficiency and success rates, and reducing labor and costs. Furthermore, flower image analysis contributes to pest and disease management by detecting early signs of plant health issues through morphological abnormalities. In this study, we established an image dataset based on the black-seeded mini watermelon cultivar, comprising 939 staminate flower images and 311 pistillate flower images. A YOLO v2 deep learning model was trained on this dataset, achieving an accuracy rate exceeding 97%. Future research will expand the database to include stem classification (main vine, secondary vine, and tertiary vine) to support the development of automated field operations for cucurbit crops.*

Keywords: *Image Recognition, Pollination Automation, Cucurbit Crops, Deep Learning.*

I. INTRODUCTION

This study focuses on cucurbit crops, such as watermelon and melon, as the primary research subjects due to their high economic value and growing market demand. Cucurbit fruits are widely consumed, especially in Asia, where their demand has been increasing annually. However, pollination management in these crops is complex, directly impacting fruit quality and yield. Cucurbit plants exhibit a monoecious flowering system with separate staminate and pistillate flowers. Farmers must manually inspect fields to ensure an optimal staminate-to-pistillate flower ratio, often removing excess flowers when necessary. Under natural pollination conditions, the success rate is only around 10%, and incomplete pollination frequently results in malformed fruits. To improve pollination rates, farmers have traditionally relied on pollinator bees or manual pollination, both of which require extensive experience and are subject to inconsistencies.

To address these challenges, we apply deep learning-based artificial intelligence (AI) to recognize staminate and pistillate flowers, transforming traditional experience-based decision-making into data-driven precision agriculture. By leveraging high-resolution imaging and feature extraction techniques, this approach enables accurate, rapid differentiation between staminate and pistillate flowers, minimizing errors and optimizing labor resources. Furthermore, precise identification of the optimal pollination timing enhances pollination success rates.

The integration of image recognition technology into the pollination process not only improves efficiency and success rates but also enhances overall crop management, increasing both yield and quality. This research aims to develop an intelligent image recognition system for distinguishing staminate and pistillate flowers in cucurbit crops, optimizing pollination strategies, and addressing key challenges in traditional agricultural practices. By automating the identification process, the system reduces human error, minimizes redundant manual operations, and significantly enhances operational efficiency and accuracy. As a result, this approach contributes to higher crop yields, reduced labor costs, and improved fruit quality, ultimately benefiting modern agricultural production.

II. LITERATURE REVIEW

To prevent excessive flowering or fruiting that could lead to nutrient competition and hinder proper fruit development, farmers often implement artificial flower and fruit thinning techniques. Excessive fruit load not only results in smaller, lower-quality fruits but also weakens plant vigor, making crops more susceptible to frost damage and diseases. Thinning flowers and fruits help regulate fruit load, adjust nutrient distribution, and optimize the leaf-to-fruit ratio. This balance significantly influences fruit growth duration and maturation. Studies have shown that modifying the leaf-to-fruit ratio affects the final fruit size and quality [1].

Cucurbit crops, including watermelon, muskmelon, cucumber, and pumpkin, exhibit distinct staminate and pistillate flowers, a characteristic that significantly impacts pollination, fruit development, and yield quality. The ability to accurately distinguish between staminate and pistillate flowers is crucial for optimizing pollination strategies, improving fruit set rates, and enhancing overall agricultural efficiency. Traditionally, farmers rely on manual observation and experience-based methods to identify and manage flowers, which can be labor-intensive, time-consuming, and prone to errors. However, advancements in computer vision and artificial intelligence (AI) are revolutionizing the precision and efficiency of flower recognition in cucurbit cultivation.

One of the primary reasons for distinguishing staminate and pistillate flowers is to ensure proper pollination. In many cucurbit species, natural pollination relies on insect activity, primarily bees, to transfer pollen from staminate to pistillate flowers. Insufficient or imbalanced pollination can lead to fruit deformation, lower yields, and poor-quality produce. In controlled environments such as greenhouses, where natural pollinators may be absent or insufficient, farmers often resort to manual pollination. Accurate flower identification allows for targeted pollination, reducing the risk of pollination failure and optimizing resource allocation.

Additionally, precise flower recognition aids in better crop management practices, such as flower thinning and pruning. By monitoring the ratio of staminate to pistillate flowers, farmers can make informed decisions to enhance plant health and fruit development. Excess staminate flowers may compete for nutrients, while an inadequate number of pistillate flowers can limit fruit production. Automated detection using AI-powered image recognition systems can provide real-time data on flower distribution, enabling farmers to adjust their management techniques dynamically.

Moreover, advanced image recognition techniques can contribute to early disease detection in flowers. Abnormalities in flower shape, color, or texture may indicate the presence of pests or diseases that could threaten the entire crop. AI-driven monitoring systems can identify these early warning signs, allowing for timely interventions that minimize damage and reduce the need for excessive pesticide use.

In conclusion, staminate and pistillate recognition in cucurbit crops is a critical aspect of modern precision agriculture. The integration of AI and imaging technologies not only enhances pollination efficiency but also improves overall crop health, reduces labor costs, and increases yield consistency. As agricultural technology continues to evolve, automated flower identification will become an indispensable tool in sustainable and high-efficiency farming.

Fen's study in 2020 explores the impact of precise nitrogen and potassium management in hydroponic systems for netted muskmelons. The research found that adjusting nutrient formulations at different growth stages improved fruit weight, dry matter ratio, and total soluble solids. These findings highlight the importance of targeted nutrient management to enhance muskmelon quality and productivity.

Musajan's research in 2024 [2] examines how digital technology influences farmers' transition to environmentally friendly muskmelon production. Using data from China's major muskmelon-growing regions, the study found that adopting digital tools significantly reduced pesticide and fertilizer use, improved precision management, and enhanced market access. The results suggest that digital agriculture can be a crucial factor in sustainable muskmelon farming.

III. MATERIALS AND METHODS

As previously mentioned, pollination is a critical factor in determining the quality and yield of cucurbit crops in field operations. Accurate monitoring of flowering conditions is essential for making informed decisions. In this study, we selected the mini watermelon variety from Known-You Seed Co., Ltd. as the primary crop. The plants were cultivated in a greenhouse environment, and a total of 1,250 images were collected (939 staminate flower images and 311 pistillate flower images) using the built-in camera of an iPhone 15.

A. *DataBase and Image Label*

In a controlled greenhouse environment, we germinated 200 mini watermelon seeds. Once the seedlings developed their second set of true leaves, they were transplanted into 10-inch soft pots, achieving an 85% seedling survival rate. Approximately 30 days post-transplantation, lateral buds below the fifth leaf were removed, while those between the fifth and tenth leaves were retained. At this stage, image collection of the flowers commenced.

Mini watermelons exhibit a monoecious flowering pattern, where staminate and pistillate flowers not only differ in morphology but also in their growth positions. Staminate flowers predominantly develop on the primary vine, while pistillate flowers appear on secondary and tertiary vines. As illustrated in Fig. 1, the blue dashed box highlights a pistillate flower, with a distinct red arrow indicating its ovary, while the yellow dashed box marks a staminate flower, which lacks the rounded ovary at its base.

For our dataset, we collected images of staminate and pistillate flowers at various blooming stages, including unopened, partially opened, and fully opened flowers. The collected images were resized to a resolution of 1920 × 1080 pixels for further processing.

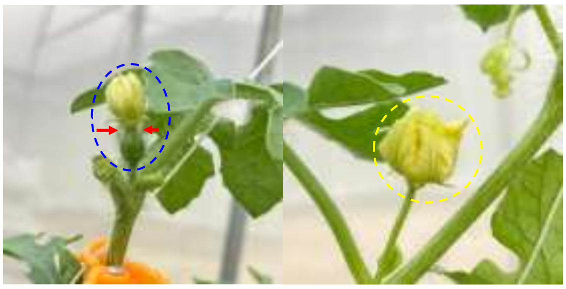

Fig. 1: Staminate (yellow dotted line) and pistillate (blue dotted line)

We used MATLAB Image Label tool, using a rectangular shape and two category labels (staminate, pistillate). For the label the staminate and pistillate, the range of the frame does not include the pedicel under the flower, and the frame includes the range of the entire flower.

B. *Deep Learning*

Deep learning encompasses various applications, including different methods for annotating ground truth data. Common annotation techniques include bounding boxes (Rectangle), pixel-wise labeling

(Pixel Label), and polygonal annotations (Polygon). The choice of annotation method influences the selection of network architectures and detection performance. For instance, bounding box annotations are widely used in object detection models such as YOLO v1–v5 [4-8] and SSD [9]. Both bounding box and polygon annotations are applicable to Mask R-CNN [10], while pixel-wise annotations are primarily utilized in semantic segmentation networks [11].

To enhance localization accuracy, this study employs a semantic segmentation network with pixel-wise annotations, using the pre-trained DeepLabv3+ [12] model. DeepLabv3+ is a convolutional neural network (CNN) specifically designed for semantic image segmentation, incorporating architectures such as Fully Convolutional Networks (FCN), SegNet [13], and U-Net [14]. Given the initial limitation in dataset size, the pre-trained model accelerates the training process, enabling preliminary performance evaluation with minimal additional data. Furthermore, the Cambridge CamVid dataset [15], which provides pixel-level annotations for 32 semantic categories, is integrated to refine object boundary detection. This integration improves edge accuracy, ensuring that detected objects align more precisely with their actual contours rather than exhibiting dispersed or inaccurate boundaries.

As discussed earlier, various AI network architectures exist for different applications. In our study, the primary objective was to detect and analyze flower counts rather than to achieve highly detailed flower shape segmentation. Therefore, we opted against using a semantic segmentation network, which requires higher hardware specifications and significantly longer computational time.

Among non-semantic segmentation networks, our target objects did not involve detecting smaller objects within larger ones. Given these considerations, we selected the YOLOv2 pre-trained model, which offers high computational efficiency and does not prioritize small-object detection, making it a suitable choice for our application.

IV. RESULT AND DISCUSSION

We implemented the YOLO v2[5] deep learning architecture with a learning rate set at 0.001 and a mini-batch size of 32. From a dataset of 1,250 images, 70% (875 images) were randomly selected as the training set, while the remaining 30% (375 images) were used as the test set. The training samples, along with their corresponding ground truth annotations, were used to optimize the network weights. To prevent overfitting and ensure the stability of detection results, the untrained test samples were utilized for validation.

To determine the optimal training epoch, we evaluated model performance at five different epochs: 200, 300, 400, 500, and 600. The model’s accuracy, precision, recall, and F1 score were computed for each case. Our results indicated that at Epoch = 400, all performance metrics exceeded 99%, with no signs of overfitting, confirming the model's stability and reliability.

*A. **Confusion Matrix***

A confusion matrix [16] is a fundamental tool for evaluating the performance of object detection models, such as YOLOv2. It assesses classification accuracy by comparing predicted results with ground truth annotations, ensuring both correct classification and proper bounding box placement. The confusion matrix consists of four key components:

True Positive (TP): Objects that are correctly classified and have an accurately placed bounding box.

True Negative (TN): Background regions or non-target objects that are correctly identified as not belonging to the target class.

False Positive (FP): Objects that are incorrectly classified or have an inaccurate bounding box, leading to false detections.

False Negative (FN): Objects that are missed by the model due to incorrect classification or improper bounding box placement.

From these values, several key performance metrics are derived:

1. *Precision (P)*: Measures the proportion of correctly detected objects out of all predicted instances. A high precision indicates fewer false positives.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

2. *Accuracy (Acc)*: Indicates the proportion of correctly classified objects out of the total dataset.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

3. *Recall (R)* / Sensitivity: Evaluates the model's ability to detect objects correctly. A high recall means fewer false negatives.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

4. *F1 Score*: The harmonic mean of precision and recall, balancing both metrics to assess overall model performance.

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

B. ***Deep Learning Recognition Rate***

To assess the performance of our object detection model, we trained YOLOv2 with EPOCH values of 200, 300, 400, 500, and 600. After training, we stored the corresponding network weights and evaluated the model on a dataset of 1250 images, analyzing the confusion matrix for each EPOCH setting, as shown in Table I.

Since each image may contain one or more objects, the evaluation criteria adjust accordingly. For example, if an image contains two objects, but only one is correctly detected and classified, the True Positive (TP) is recorded as 0.5, and the True Negative (TN) is also 0.5 to reflect partial detection accuracy.

At Epoch = 200 and 300, the detection results were inconsistent, leading to variations in TP and TN values. Notably, at Epoch = 200, TP and TN were higher than at Epoch = 300, indicating fluctuations in detection stability. As training progressed to Epoch = 400, the model's recognition ability improved significantly, achieving a stable detection rate.

Beyond Epoch = 400, the model successfully classified and detected at least 1210 out of 1250 images correctly, with TN and FP values remaining below 50, ensuring a high detection reliability.

To enhance localization accuracy, this study employs a semantic segmentation network with pixel-wise annotations, using the pre-trained DeepLabv3+ [12] mod

Table I: Confusion Matrix Analysis For Different Epoch Settings.

Epoch	TP	TN	FP	FN
200	1198	21	31	0
300	1188	43	19	0
400	1210	30	10	0
500	1217	27	6	0
600	1231.7	18.3	0	0

C. *Stability Analysis*

To further evaluate the stability of the trained network weights, we conducted a detailed analysis using the confusion matrix at different EPOCH settings. By applying Equations (1)–(4), we assessed key performance metrics, including Precision, Accuracy, Recall, and F1 Score, as shown in Table II.

The stability of the network weights was determined by monitoring fluctuations in these metrics across different EPOCH values. A stable model should maintain consistently high values for Precision, Accuracy, Recall, and F1 Score, while minimizing variations in False Positives (FP) and False Negatives (FN).

Through our evaluation, we observed that as EPOCH increased beyond 400, the detection performance reached an optimal and stable state, with all four metrics exceeding 99%. The consistency of these results confirms that the network weights effectively generalize to unseen data, ensuring reliable object detection.

Table III: Performance Metrics Derived From The Confusion Matrix Using Equations (1)–(4).

Epoch	200	300	400	500	600
Accuracy	97.52	98.48	99.2	99.52	100
Precision	97.47	98.42	99.18	99.5	100
Recall	100	100	100	100	100
F1 score	98.71	99.2	99.58	99.74	100

Author details must not show any professional title (e.g. Managing Director), any academic title (e.g. Dr.) or any membership of any professional organization (e.g. Senior Member ONGC).

To avoid confusion, the family name must be written as the last part of each author name (e.g. John A.K. Smith).

Each affiliation must include, at the very least, the name of the company and the name of the country where the author is based (e.g. Causal Productions Pty Ltd, Australia).

Email address is compulsory for the corresponding author.

V. CONCLUSION

Accurate identification of staminate and pistillate in cucurbit crops, such as watermelon and muskmelon, is crucial for optimizing pollination, fruit development, and yield quality. Traditional manual identification is time-consuming and error-prone, while AI-powered image recognition offers precision and efficiency. Proper pollination management ensures high fruit set rates, preventing deformities and low yields. In greenhouses, where natural pollinators are scarce, targeted manual pollination based on accurate flower detection improves productivity. Additionally, monitoring flower ratios helps optimize pruning and nutrient allocation, enhancing plant health. AI-driven detection also enables early disease identification by recognizing abnormalities in flower shape or color. Integrating advanced imaging technology in agriculture reduces labor costs, increases efficiency, and improves yield quality, making it a vital tool for modern precision farming.

This study utilized a dataset of 1,250 images, consisting of 939 staminate images and 311 pistillate images. The dataset was randomly split, with 70% used for training and 30% for testing. Using a YOLO v2 pre-trained network, we fine-tuned the model by integrating our dataset and training it for 500 iterations with a batch size of 32 and a learning rate of 0.001. The retrained model achieved an accuracy of 99.52%, a precision of 99.5%, a recall of 99.5%, and an F1 score of 99.58%. Future work will expand the dataset to 3,000 images and incorporate images of primary, secondary, and tertiary vines. These additional categories will be included in the training process to further advance the automation of field operations for cucurbit crop management.

ACKNOWLEDGEMENT

This work was supported by the National Science and Technology Council (NSTC) of Taiwan (NSTC 113-2313-B-020-010). We thank to National Center for High-performance Computing (NCHC) of National Applied Research Laboratories (NARLabs) in Taiwan for providing computational and storage resources.

AUTHOR CONTRIBUTIONS

Kuo-Dung Chiou, Yu-Shen Liang and Chia-Ying Chang conducted the research; Yi-Zhen Chen, Shin-Hau Chiou analyzed the data; Kuo-Dung Chiou and Chia-Ying Chang wrote the paper and program coding; all authors have approved the final version.

REFERENCES

[1] I-Hsuan Hung, "Sink Change Affects Vegetative and Generative Growth of. Papaya." Diss. 2011.

[2] Fan, Yen-Tzu, et al. "Quality Improvement of Netted Melon (Cucumis melo L. var. reticulatus) through Precise Nitrogen and Potassium Management in a Hydroponic System." Agronomy, vol. 10, no. 6, p. 816, 2020. https://doi.org/10.3390/agronomy10060816.

[3] Musajan, Adiljan, et al. "Unveiling the Mechanisms of Digital Technology in Driving Farmers' Green Production Transformation: Evidence from China's Watermelon and Muskmelon Sector." Foods, vol. 13, no. 23, p. 3926, 2024. https://doi.org/10.3390/foods13233926.

[4] Joseph Redmon, et al. "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788, 2016.

[5] Joseph Redmon, and Ali Farhadi, "YOLO9000: Better, Faster, Stronger," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7263-7271, 2017.

[6] Joseph Redmon, and Ali Farhadi, "YOLOv3: An Incremental Improvement," Tech report, 2018.

[7] Alexey Bochkovskiy, et al. "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv 2020, arXiv:2004.10934.

[8] Xingkui Zhu, et al. "TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios," arXiv:2108.11539.

[9] W. Liu, et al. "SSD: Single shot multibox detector," Computer Vision – ECCV, vol. 9905, 2016.

[10] He, Kaiming, et al. "Mask R-CNN." ArXiv:1703.06870 [Cs], January 24, 2018.

[11] R. Kemker, et al. "High-Resolution Multispectral Dataset for Semantic Segmentation." CoRR, abs/1703.01918. 2017.

[12] Liang-Chieh Chen, et al. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," ECCV, 2018.

[13] V. Badrinarayanan, et al. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 12, 2481-2495, 2017.

[14] O. Ronneberger, et al. "U-Net: Convolutional Networks for Biomedical Image Segmentation," Lecture Notes in Computer Science, vol 9351. 2015.

[15] G. J. Brostow, et al. "Semantic object classes in video: A high-definition ground truth database," Pattern Recognition Letters, vol. 30, no. 2, 88-97, 2009.

[16] J. Liang, "Confusion matrix: machine learning, " POGIL Activity Clgh., vol. 3, no. 4, 2022.