Computers and Electronics in Agriculture 201 (2022) 107342



Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag





Multi-class detection of kiwifruit flower and its distribution identification in orchard based on YOLOv5l and Euclidean distance

Guo Li [a], Longsheng Fu [a,b,c,d,*], Changqing Gao [a], Wentai Fang [a], Guanao Zhao [a], Fuxi Shi [a], Jaspreet Dhupia [e], Kegang Zhao [f], Rui Li [a,d], Yongjie Cui [a]

[a] College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
[b] Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China
[c] Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China
[d] Northwest A&F University Shenzhen Research Institute, Shenzhen, Guangdong 518000, China
[e] Department of Mechanical Engineering, The University of Auckland, Private bag, Auckland 92019, New Zealand
[f] Xi'an Agriculture Machinery Management and Extension Station, Xi'an 710065, China

ARTICLE INFO

Keywords:
Kiwifruit flower
YOLOv5l
Flower phenology
Flower distribution
Robotic pollination

ABSTRACT

Asynchrony of kiwifruit flowering time results in different flower phenological stages in canopy at the same time. Pollination quality of flowers is influenced by their phenological stages, while their distributions determine fruit distributions and influence kiwifruit quality and yield. Thus, it's necessary to find suitable flowers to be pollinated based on flower phenology and its distribution. However, influences of flower phenology and flower distribution were not considered in most previous studies about robotic pollination of kiwifruit, where pollination of all open flowers was indiscriminate. Therefore, a method was proposed for multi-class detection of kiwifruit flower and its distribution identification in orchard, which was based on You Only Look Once version 5 large (YOLOv5l) and Euclidean distance. According to kiwifruit flower phenology, kiwifruit flowers were classified into 10 classes to find suitable flowers for pollination, while flower cluster and branch junction were divided into 4 classes for obtaining flower distributions. All classes were manually labeled by rectangular bounding boxes for training and testing. Considering high detection accuracy requirements with small model size, YOLOv5l was applied to do transfer learning for multi-class detection of kiwifruit flower. Then, pixels coordinate of multi-class objects and their corresponding Euclidean distances could be gained. Finally, flower distributions in canopy were obtained by matching method. Total mean Average Precision (mAP) was 91.60 % in YOLOv5l, while the mAP of multi-class flower (10 classes) was 93.23 %, which was 5.70 % higher than that of the other 4 classes. Matching accuracy (MA) of flowers matching flower clusters was up to 97.60 %. Moreover, MA of flower cluster matching branch junction (96.20 %) and total MA (93.30 %) increased by 1.20 % and 1.00 % based on improved matching method, respectively. Total processing speed of multi-class flower detection and its distribution identification was 112.46 ms per image including 15.50 ms for image detection by YOLOv5l. Results showed that multi-class kiwifruit flowers and relative flower distributions could be fast and accurately obtained for further selecting suitable flowers for robotic pollination.

1. Introduction

Pollination in suitable flower phenology is critical to achieving satisfactory kiwifruit production and quality. Kiwifruit size is directly related to seed number, i.e., to the number of fertilized ovaries, which depend on pollination (Gonzalez et al., 1995). However, kiwifruit is more difficult than other species on pollination, since not only it's dioecious without synchrony of flowering time, but also its flowers lacking nectar are not enough attractive to pollinators (Castro et al., 2021). Artificial pollination overcame difficulty and became a critical technique to increase kiwifruit quality (Lim et al., 2020). Nevertheless, artificial pollination does not always reach maximal efficiency, which depends on female flower phenological stages at pollination, a pollination system (dry or liquid), and pollination methods (Gianni and Vania, 2018). Flower phenology has been considered during artificial pollination by workers, which was hardly to implement automatically in

* Corresponding author at: College of Mechanical and Electronic Engineering, Northwest A&F University, China.
E-mail address: fulsh@nwafu.edu.cn (L. Fu).

https://doi.org/10.1016/j.compag.2022.107342
Received 16 June 2022; Received in revised form 31 July 2022; Accepted 24 August 2022
Available online 2 September 2022
0168-1699/© 2022 Elsevier B.V. All rights reserved.

robotic pollination. Therefore, it is necessary to develop robotic pollination based on flower phenology, which can not only replace manual pollination to save high labor costs but also improve pollination quality of kiwifruit flowers to increase kiwifruit quality and size.

Unbalanced flower distributions may cause unbalanced crop load and nutrient delivery, which can be improved by flower and fruit thinning. Early flowers (usually in middle part of fruit branch) had larger ovaries with more locules and ovules than late flowers (in base portion or top of fruit branch) on the same vine and they produced larger fruit (McPherson et al., 2001). Comparisons between early and late flowers may be a result of competition for resources within fruit branch and positional effects (Smith et al., 1994). Although flower thinning or fruit thinning was usually taken to adjust flower or fruit distributions (fruit distributions mainly depending on flower distributions), flower thinning was more effective in enhancing fruit size and weight as compared to fruitlet thinning (Cangi and Atalay, 2018; Wu et al., 2022). Besides, without pollination or invalid pollination, ovary of kiwifruit flower had drop phenomenon (Thakur and Chandel, 2004). Therefore, some suitable flowers (mainly in middle part of fruit branch) could be chosen to be pollinated based on their distributions to replace flower thinning and decrease work of fruit thinning.

Kiwifruit flowers can be divided into multi-class flowers based on their flower phenology to select suitable flowers for pollination. Researchers did many relative studies on classification of kiwifruit flowers. Salinero et al. (2009) proposed BBCH (named after its developing institutes Biologische Bundesanstalt, Bundessortenamt, und CHemische Industrie) scale of kiwifruit plant, which divided flowering phenology into closed flower, white petals, ocher petals, early petal fall, and petal fall. Although the BBCH scale of kiwifruit plant provided information on flower phenology, relations between phenological flowering stages and pollination were vague. According to phenological flowering stages of kiwifruit, Gianni and Vania (2018) divided flowers into 5 classes before pollination, i.e., closed flower, white petals, ocher petals, early petals fall, and petals fall. They found out that dry pollination was better made with pure pollen before petal fall stage, while liquid pollination was not later than early petal fall stage. Previous studies about classification of kiwifruit flowers were just used for optimal orchard management such as artificial pollination and flower thinning (mainly relying on humans). However, with rapid development of agricultural robot, it's necessary to combine knowledge of agronomic pollination with robot for precision pollination. Thus, multi-classification of kiwifruit flowers was essential for robotic pollination based on flower phenology.

It is very challenging to identify flower phenology and flower distributions due to complex kiwifruit orchard environments (such as overlapping flowering phenology and occluding situations), which may be solved by image processing algorithms. Most previous studies focused on identification of multi-class fruit for harvesting or multi-class flower for flower thinning, where image processing algorithms were mainly divided into traditional image processing technology and deep learning methods. Traditional approaches mainly focused on features such as color and texture using complex algorithms with many fixed thresholds, which probably produced non-robust and inaccurate results in complex environments of orchard (Fu et al., 2021; Tian et al., 2020; Majeed et al., 2020). With rapid development of deep learning methods, a considerable number of researchers applied them for classification tasks in agricultural field, which conquered limitations of traditional image processing technology (Dias et al., 2018; Wu et al., 2021; Zhang et al., 2020; Zhang et al., 2021). Gao et al. (2020) proposed a multi-class apple (non-occluded, leaf-occluded, branch/wire-occluded, and fruit-occluded fruit) detection method based on Faster Region-Convolutional Neural Network (R-CNN) with Visual Geometry Group with 16 layers (VGG16), which obtained mean Average Precision (mAP) of 87.85 % and detection speed of 241.0 ms per image with 1920 × 1080 pixels. Although Faster R-CNN (classic of two-stage networks) achieved high accuracy in complex orchard environments, its detection speed was slow. Therefore, object detection models need to be selected in terms of both accuracy and speed.

You Only Look Once (YOLO), as a one-stage network, showed high potential in speed while achieving high accuracy in agricultural field. Suo et al. (2021) employed YOLOv4 for identification of five-classes kiwifruit (fruit not occluded, fruit occluded by leaves, fruit occluded by other fruits, fruit occluded by branches, fruit occluded by wires), which achieved mAP of 91.9 % and detection speed of 25.5 ms per image with 2,352 × 1,568 pixels. Jocher et al. (2020) proposed YOLOv5 in PyTorch framework, the latest version of YOLO family, which was divided into many versions based on model complexity, namely YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, and so on. Therefore, compared with YOLOv3 and YOLOv4, YOLOv5 had a more flexible network structure to choose appropriate one more conveniently based on different tasks (Bochkovskiy et al., 2020; Kuznetsova et al., 2020; Lin et al., 2021; Redmon and Farhadi, 2018). Then, in another field of agriculture, Wang et al. (2021) proposed a method for estimation of apple flower phenology distribution based on YOLOv5l to gain timing of flower thinning, which acquired classification accuracy up to 92 % and inference time of 16 ms per image with 4,196 × 2,160 pixels. YOLO models, especially YOLOv5, achieved considerably satisfactory results on classification tasks in orchard. Therefore, among YOLOv5 models, YOLOv5l was selected for this study since this model attained high detection accuracy and fast detection speed.

This study aimed to develop a multi-class detection of kiwifruit flower and its distribution identification method based on YOLOv5l and Euclidean distance to select suitable kiwifruit flowers for robotic pollination. This work is organized as follows. Section 2 will describe the design of this proposed method; Section 3 will present the obtained results of multi-class kiwifruit flower detection and matching among multi-class objects in canopy, and discuss the relevant issues; Section4 will present the conclusions and prospects acquired from this study.

2. Materials and methods

As mentioned, this study has focused on developing an image processing system based on YOLOv5l and Euclidean distance to select suitable flowers for robotic pollination in kiwifruit orchard. Firstly, original images were augmented and then manually labeled (using rectangular bounding boxes to label multi-class flowers based on kiwifruit flower phenology such as bud, half-open, fresh pistil, ocher pistil, petal fall, and other objects including flower clusters, occluding flowers, branch junctions) for training and testing. YOLOv5l was trained based on its pre-trained weights, and then it was applied for detection of above-mentioned multi-class objects. Then, multi-class flowers matched relative flower clusters and branch junctions, after calculating Euclidean distances between flower clusters and branch junctions. Finally, AP, mAP, confusion matrix, average detection speed, and matching accuracy (MA) were employed to evaluate the proposed methodology.

2.1. Image acquisition

Images of "Xuxiang" (Actinidia deliciosa) kiwifruit flowers were acquired during the flowering season of 2021 from Meixian Kiwifruit Experimental Station in Shaanxi province, China. Kiwifruits are commercially grown on sturdy support structures such as T-bars and pergolas resulting in that kiwifruit canopy is one layer of kiwifruit vine (Song et al., 2021). Kiwifruit flower maybe adjacent or overlapped in canopy, but the pistil of each flower is generally visible and independent. "Xuxiang" is widely planted in this area, and most of its flowers are hanging down from canopy and facing to ground. An ordinary single-lens reflex camera (Canon S110, Canon Inc., Tokyo, Japan) on "AUTO" mode with a resolution of 3,456 × 2,304 pixels was held at around 50 cm below flowers to shoot them upwards. A total of 355 original images were acquired at different natural lighting conditions, which were saved in Portable Network Graphics (PNG) format. Some examples of different lighting conditions were shown in Fig. 1.

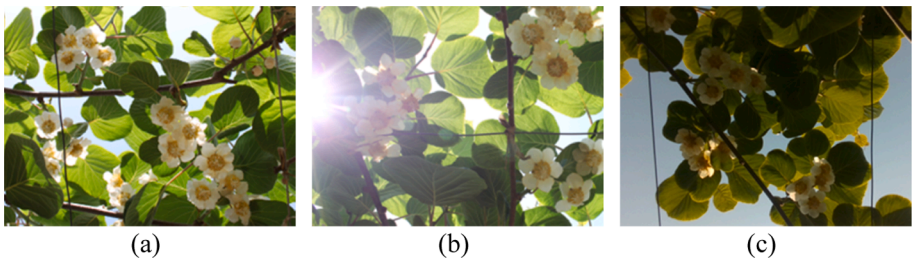


Fig. 1. Examples of kiwifruit flower and bud images acquired in orchard. (a), (b), and (c) represented front-lighting condition, glare condition caused by sunlight, and back-lighting condition, respectively.

2.2. Image dataset

Kiwifruit flowers were labeled into multiple classes based on flower phenology to acquire optimal pollination timing. Those images were manually annotated by using a software tool, LabelImg (https://github.com/tzutalin/labelImg), where rectangular bounding boxes were drawn to label multi-class flowers, i.e., Bud (BD), Early open (EO), Half-open (HO), Fresh pistil (FP), Early ocher pistil (EOP), Ocher pistil (OP), and Petal fall (PF), with their acronyms as labels. Labeling examples of multi-class kiwifruit flowers with different colors were shown in Fig. 2. Only pistil of flower received pollen and produced seeds, and it was thus labeled to help robot avoid spraying pollen in other areas of flower. Examples of labeling only pistil were shown in Fig. 2d to Fig. 2f. If a pistil was not visible or partly visible, its whole flower would be labeled by a rectangular bounding box, as shown in Fig. 2a to Fig. 2c. In addition, if one or more petals fell, due to the whole part possessing more obvious features without confusion, it was labeled instead of pistil part, as shown in Fig. 2g.

Some flowers were labeled into three new classes due to features of being occluded or being sprayed with colored pollen. If pistil was occluded by other objects (leave or flower), pollen could hardly be sprayed in it. Thus, occluded pistil of flower (OF) was labeled as a new class, as shown in Fig. 2h. If one flower was pollinated, its pistil was partly stained with colored pollen (usually pink), which was used to differentiate between pollinated and unpollinated flowers. Therefore, pistils with colored pollen were labeled into another two classes ("Bright pollen (BP)" and "Dark pollen (DP)") based on the time of spraying pollen, as shown in Fig. 2i to Fig. 2j, where the BP represented that pistil was just pollinated, while the DP represented that pistil was pollinated for a few days. Although three classes in Fig. 2h to Fig. 2j did not belong to kiwifruit flowering phenological stages, they were still divided into classes of multi-class flowers according to their relationship with pollination and their features in this study.

In addition, flower clusters and branch junctions were also labeled, which were the basis for getting flower distributions. White-yellow flowers were produced singly or in clusters at leaf axils of kiwifruit

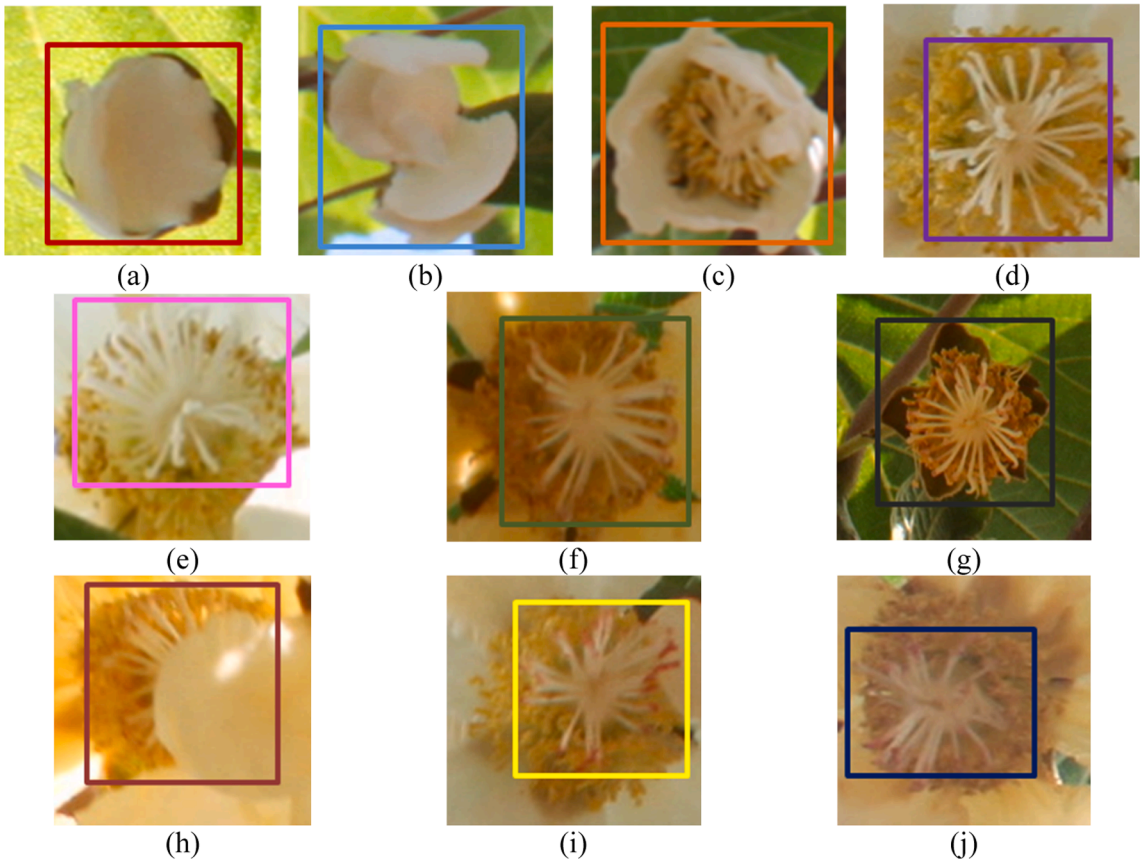


Fig. 2. Labeling examples of multi-class kiwifruit flowers with different colors. (a) Bud (BD); (b) Early open (EO); (c) Half-open (HO); (d) Fresh pistil (FP); (e) Early ocher pistil (EOP); (f) Ocher pistil (OP); (g) and Petal fall (PF); (h) Occluded pistil of flower (OF); (i) Bright pollen (BP); (j) Dark pollen (DP).

branch, and junction of kiwifruit branch and main branch was branch junction. Thus, it's necessary to detect flower clusters and branch junctions for further obtaining flower distributions. Objects classified into more classes based on their obvious features are promising to achieve better detection results (Suo et al., 2021). Therefore, branch junctions were classified into three classes based on occlusion situation, as shown in Fig. 3, where an example of labeling a flower cluster was also shown. After annotation per image in LabelImg tool, annotation files were saved in "xml" format, and then they were converted to specific annotation files in "txt" format of YOLO.

After being labeled, image dataset was divided and augmented. All original images (355 images) were randomly divided into a training set (284 images) and a validation set (71 images) with a ratio of 4 to 1. Then, the training set was augmented to 1704 images with image mirroring (including horizontal and vertical mirroring) and image rotation (including 90°, 180°, and 270° of rotation) to improve generalization ability of deep learning model in this study. The original and augmented images datasets were available on https://github.com/fu3lab/Kiwifruit_multi-class_flower_images.

2.3. Network architecture of YOLOv5

YOLOv5, the latest version of YOLO family in 2021, showed high potential in speed while achieving high accuracy. Cross Stage Partial Network (CSPNet), as backbone of YOLOv5, extracts image features by merging cross-stage hierarchical structure, which reduces the repeated gradient information and floating-point operations per second (FLOPS). Before forwarding to feature aggregation architecture in neck, output feature maps of CSPNet backbone were sent to Spatial Pyramid Pooling block (SPP), which increases receptive field by converting any size of feature map into a fixed-size feature vector (Wang et al., 2020b). YOLOv5 adopts Path Aggregation Network (PANet) structure as its neck based on Feature Pyramid Networks (FPN) structure, which enhances bottom-up path and improves propagation of underlying features in process of extracting features (Lin et al., 2017; Liu et al., 2018). Besides, this structure not only enriches scales of output feature map but also improves extraction ability of target features by combining different layer features (Wang et al., 2020a; Xu et al., 2021). Finally, head in YOLOv5, namely YOLO layer, is the same as YOLOv4 and YOLOv3, which generates three different scales of feature maps to achieve multi-scale prediction, enhancing prediction of small, medium, and large objects in model (Nepal and Eslamiat, 2022). YOLOv5′s network architecture using default network parameters is shown in Fig. 4.

Compared with YOLOv3 and YOLOv4, YOLOv5 has a more flexible network structure, whose parameters can be adjusted for different tasks conveniently. YOLOv5 models were divided into many versions based on model complexity, namely YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, and so on. From YOLOv5s to YOLOv5x, although accuracy will be improved with the increase in model complexity (model depth and layer channel), detection speed and model size will decrease. Due to appropriate balance of detection accuracy, detection speed and model size, YOLOv5l was applied as the detection model in this study.

2.4. Network training

Experiments were carried out based on deep learning framework of PyTorch platform with version 1.9.1 on a computer with AMD Ryzen 7 5800X 8-Core Processor (3.80 GHz) CPU, NVidia GeForce GTX 3080 Ti 12 GB GPU (10,240 CUDA cores), and 64 GB of memory, running on a Windows 10 64-bit system. Software tools included CUDA 11.1, CUDNN 8.0.5, OpenCV 4.1.2. YOLOv5l was used for training multi-class objects detection network by transfer learning based on PyTorch framework. Transfer learning is a machine learning method, which refers to a pre-trained model being reused for training in another task. Some parameters of training YOLOv5l, i.e., input size, batch size, learning rate, momentum, decay, epoch, and class number, were could be seen in Table 1.

2.5. Multi-class detection of flower and its distribution

YOLOv5l was used to achieve multi-class training and detection of objects (i.e., multi-class kiwifruit flowers, flower clusters, and branch junctions). The pipeline of labeling, training and detection was shown in Fig. 5a to Fig. 5d. Flowers in suitable pollination timing could be found out from detection results by YOLOv5l. Besides, flower clusters and branch junctions were detected by YOLOv5l to further calculate flower distributions. Detection examples of flowers, flower clusters and branch junctions were shown in Fig. 5d.

Flower distributions were obtained by matching method based on Euclidean distance. Each fruit branch or its flower cluster is separated by a suitable distance in commercial orchard, where each flower cluster usually has only one closest and corresponding branch junction. Firstly, according to detection results of YOLOv5l, both detection boxes of multi-class objects and center pixel coordinates of flower clusters were obtained. Secondly, detection boxes of flower clusters were regarded as the location threshold, and center pixel coordinates of multi-class flowers mapped those detection boxes. If a center pixel coordinate of flower was in a detection box of flower cluster, the flower matched the flower cluster, as shown in Fig. 5e. Then, Euclidean distances between every flower cluster and all branch junctions were calculated to gain correct matching results between flower clusters and branch junctions. Each flower cluster had a minimum Euclidean distance that was used to find its corresponding branch junction. If the corresponding branch junction was invisible, the minimum Euclidean distance was greater than a threshold (width or length of the flower cluster's detection box) resulting in matching no corresponding branch junction. Examples of matching results (namely flower distributions) among multi-class flowers, flower clusters and branch junctions, were shown in Fig. 5f.

Original match method might obtain small number of incorrect matching results. Flower cluster might match a closest branch junction instead of its corresponding branch junction. Thus, an improved matching method was proposed based on the minimum Euclidean distances of global matching results. If two or more flower clusters matched a same branch junction, a flower cluster was selected to match the branch junction based on the minimum Euclidean distances. In addition, effective minimum Euclidean distances should be less than the same threshold as the original matching method. Then, the rest flower clusters



Fig. 3. Other labeling objects in canopy. (a) Flower cluster (FC); (b) Branch junction without being occluded (BJ); (c) Partially occluded branch junction (OJ); (d) Completely occluded branch junction (OB).

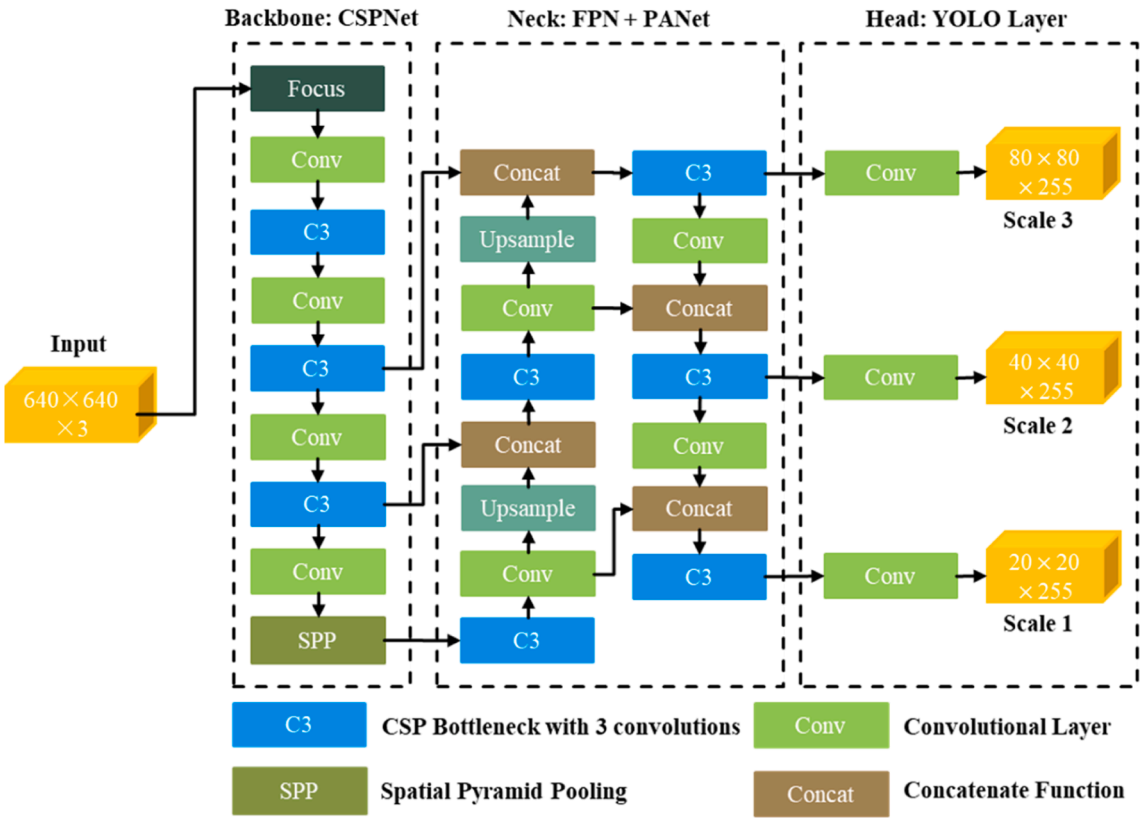


Fig. 4. The network architecture of YOLOv5. It consists of four parts, i.e., (1) Input: image with 640 × 640, (2) Backbone: CSPNet, (3) Neck: FPN + PANet, and (4) Head: YOLO Layer. Focus layer is spatial to channel layer that warps (H × W × 3) into (H/2 × W/2 × 12). Data are input to CSPNet for feature extraction at first and then fed to PANet for feature fusion. Finally, YOLO Layer generates three different outputs of feature maps to achieve multi-scale prediction.

Table 1
Network parameters of YOLOv5l.

Model	Input	Batch size	Learning rate	Momentum	Decay	Epoch	Class number
YOLOv5l	640 × 640	16	0.01	0.937	0.0005	300	14

rematched the rest branch junctions until matching over.

2.6. Performance evaluation

Loss curve, AP, mAP, confusion matrix, and detection speed (evaluation indicators of network) were used to evaluate trained model. Loss curves were employed to tell whether YOLOv5l was convergent and overfitted or not during training. AP_k was defined as the area under P_k - R_k curve (the P_k as the vertical axis and the R_k as the horizontal axis) in Eq. (1), which was used to evaluate performance of model in detecting each class. The value k represented each class of objects in this study. The mAP was defined in Eq. (2) as the average AP of n classes. The higher AP and mAP are, the better detection results of deep learning model acquire for a given object. Also, confusion matrix and detection speed were calculated to evaluate performance of YOLOv5l.

MA is one metric used to evaluate matching method for flower distribution, as defined in Eq. (3). Taking flower cluster matching branch junction as an example, if one flower cluster corresponds to one branch junction and matches it, this is true matching positive (TMP); if one flower cluster corresponds to one branch junction but matches another branch junction, this is false matching negative (FMN); if one flower cluster does not have corresponding branch junction but matches one branch junction, this is false matching positive (FMP). Although MA depended on detection results, if flowers (or branch junctions) were only incorrectly detected as other multi-class flowers (or branch junctions), it would not change matching results. MA of flowers matching flower cluster and MA of flower cluster matching branch junction were represented MA_{fmc} and MA_{cmj}, respectively. In addition, MA of consecutive twice matching (namely flowers matching a relative flower cluster and then the flower cluster matching branch junction) was represented by total MA (MA_t). Besides, total processing speed of detection and object matching was calculated to tell whether it met requirements of robotic pollination.

$$AP_k = \int_0^1 P_k(R_k)\, dR_k \tag{1}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \tag{2}$$

$$MA = \frac{TMP}{TMP + TMN + FMP} \tag{3}$$

3. Results and discussion

3.1. Training evaluation

Loss was defined as a measure of how far a model's predictions were from its label. Training loss curves of YOLOv5l had been converging, as shown in Fig. 6, where abscissa and ordinate represented training epoch

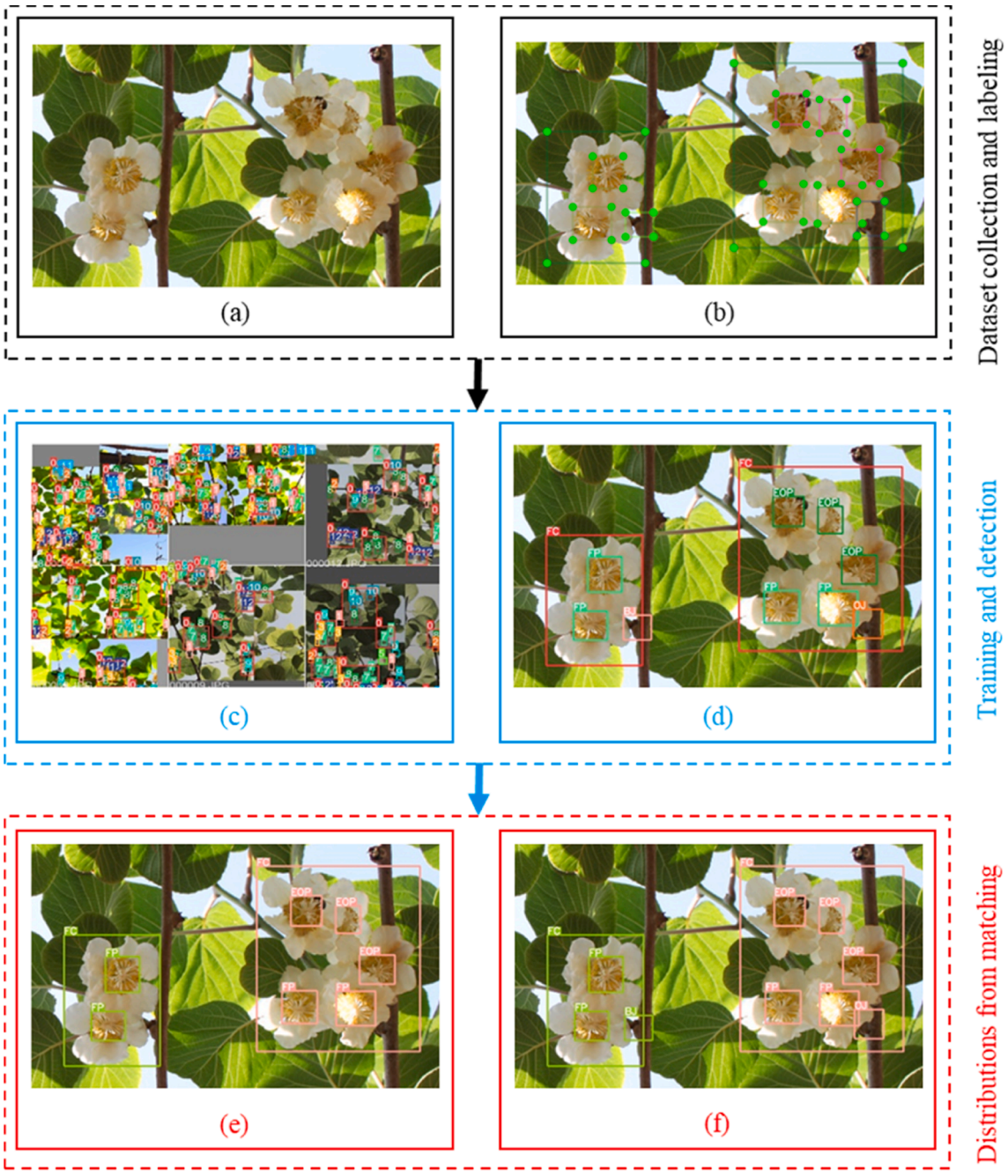


Fig. 5. The pipeline of acquiring multi-class detection of flowers and their distributions. Dark dotted rectangles showed dataset collection and labeling; Blue dotted rectangles showed training and detection of YOLOv5l; Red dotted rectangles showed obtaining flower distributions by matching method. (a) Dataset collected in kiwifruit orchard; (b) Labeling with bounding rectangles; (c) YOLOv5l training with mosaic data augmentation; (d) YOLOv5l detected multi-class flowers, flower clusters, and branch junctions with different colorful rectangles and labels. (e) Results of flowers matching flower cluster, where the same color rectangles represented matching result per flower cluster; (f) Results of flower clusters matching branch junctions, where color of branch junction was same as its relative flower cluster. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and loss values. As the number of training epochs continually increased, the loss values of YOLOv5l decreased fast at first and then decreased slowly to convergence. There were six loss curves of YOLOv5l shown in Fig. 6, where different colors and line types represented different losses in train or validation (val). The box loss (box_loss) was localization loss and loss of bounding box regressor. The objectness loss (obj-loss) was caused by an incorrect box-object IoU prediction, where a predicted bounding box was an object of background. The classification loss (cls-loss) indicated loss for classification of detected objects into various classes. The box_loss curve of YOLOv5l gradually converged near 0.1 after approximately 270 epochs, while obj-loss curve and cls-loss curve of YOLOv5l converged near 0.022 and 0.05 after approximately 200 epochs, respectively. Besides, as shown in Fig. 6, three val loss curves decreased with epochs increasing, which meant that the YOLOv5l network was not overfitted. Loss curves of YOLOv5l converged to very low values, which demonstrated that trained YOLOv5l had learned well features of 14 classes with believable predicted outputs.

3.2. Results and evaluation of YOLOv5l

Most performance indices, including AP, mAP, confusion matrix, and average detection speed, were used to verify YOLOv5l. The total mAP

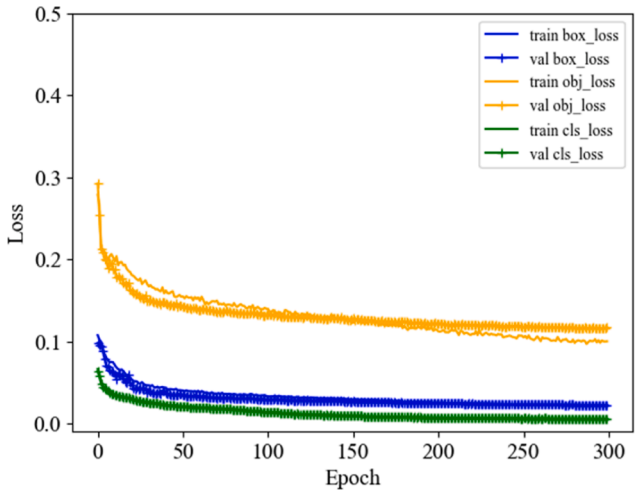


Fig. 6. Loss curves of YOLOv5l. Blue, green, and orange curves represented box-loss, obj-loss curve, and cls-loss curves of YOLOv5l, respectively. Different line types represented loss in train or val. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

was 91.60 %, while the mAP of multi-class flower (10 classes) by YOLOv5l was 93.23 %, which was 5.70 % higher than that of other objects (4 classes), as shown in Fig. 7. The reason was that objects of multi-class flower (10 classes) were much more than objects of other objects (4 classes), which meant that YOLOv5l learned more target features of multi-class flowers and reached better mAP.

Although high mAP of multi-class flowers (10 classes) was achieved, APs of some classes (EO and OF) were below 90.00 %, which could be analyzed by their confusion matrix. Fig. 8 showed a confusion matrix of 14 classes by YOLOv5l at a confidence threshold of 0.25. According to Fig. 8, proportions of false negative (FN) only existed on EOP, OF, FC, and background, which were much low with 0.02, 0.01, 0.02, and 0.02, respectively. The reason was that feature of FP was more obvious than other classes as shown in Fig. 2. On the contrary, features of some objects on EO and OF might be similar to other classes resulting in low APs, especially on EO. As shown in Fig. 8, prediction proportion was only 0.67, while proportions of FN were 0.13 on BD, 0.10 on HO, and 0.10 on background, respectively. Considering Fig. 2, features of EO might be a little similar to that of BD. Significantly, although there was a certain correlation between confusion matrix and AP, they were not equivalent to each other because of both different definitions and calculation methods. From the above results, despite some classification shortcomings in EO and OF, YOLOv5l still achieved satisfactory results on classification tasks of flowers.

The APs of other objects (4 classes) could be also analyzed by their confusion matrix. As shown in Fig. 7, APs of FC and BJ detection were all more than 94.00 % while those of OJ and OB were below 90.00 %, especially on OB with 71.20 %. Features of FC and that of BJ were more different than other classes, as shown in Fig. 3, there was thus little confusion with other classes. According to Fig. 8, there were some proportions of FN with low values on FC, while prediction on BJ obtained only one proportion of FN, which was 0.06 on background. The above contents explained why there were such high APs on FC and BJ. As for OJ and OB, they were just distinguished by situation of occlusion (partially or completely), where there might be confusion between them. As shown in Fig. 8, the confusion between OJ and OB was not necessarily between classes, but more likely between detections and backgrounds. Confusion in OJ prediction was mainly in BJ with the proportion of 0.13, while that in OB prediction was mainly in background, whose FN proportion was up to 0.24. OJ was just partially occluded by petals or leaves, and thus its features were a little similar to BJ, as shown in Fig. 3. Moreover, OB was almost completely occluded by petals or leaves causing its area within the rectangular bounding box to be more like background. Therefore, APs of both OJ and OB were not high and resulted in only 87.53 % of mAP on other objects.

Although there were no reports on detection of multi-class kiwifruit flower based on flower phenology in previous robotic pollination studies, it was reported that kiwifruit flowers were classified into one or two classes (i.e., flower or flower and bud) to meet the most basic requirement of robotic pollination. As shown in Table 2, Williams et al. (2020) obtained mAP of 85.3 % on kiwifruit flower detection by Faster R-CNN with model size of 142 MB, where not only mAP of 7.93 % was lower than that of our study, but also the model size of 52.6 MB was larger than that of YOLOv5l. Besides, detection speed of Faster R-CNN was 100.00 ms per image at 1024 × 600 image pixels, which was 84.50 ms slower than YOLOv5l at 4608 × 3456 image pixels. Li et al. (2022) applied YOLOv4 to train and detect kiwifruit flower and bud only with 38.64 ms of detection speed and model size of 244 MB, which obtained APs of flower and bud detection of 96.66 % and 98.57 %, respectively. As shown in Table 2, although APs of flower and bud by YOLOv4 were 3.54 % and 4.37 % higher than these of YOLOv5l, respectively, not only detection speed of YOLOv4 was 23.14 ms slower than that of YOLOv5l, but also model size of YOLOv4 was 156.4 MB larger than that of YOLOv5l. From the above results, YOLOv5l attained satisfactory

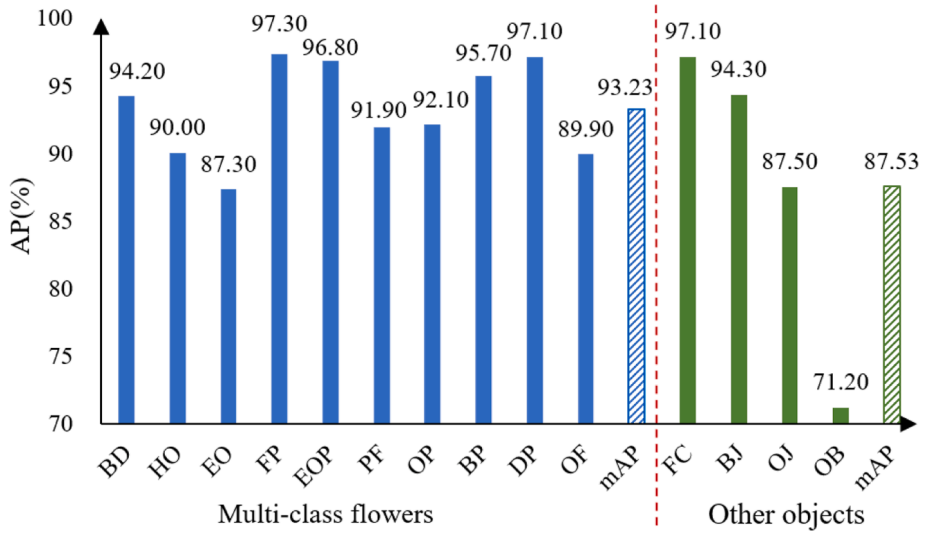


Fig. 7. Detection results of multi-class flowers and other objects with YOLOv5l.

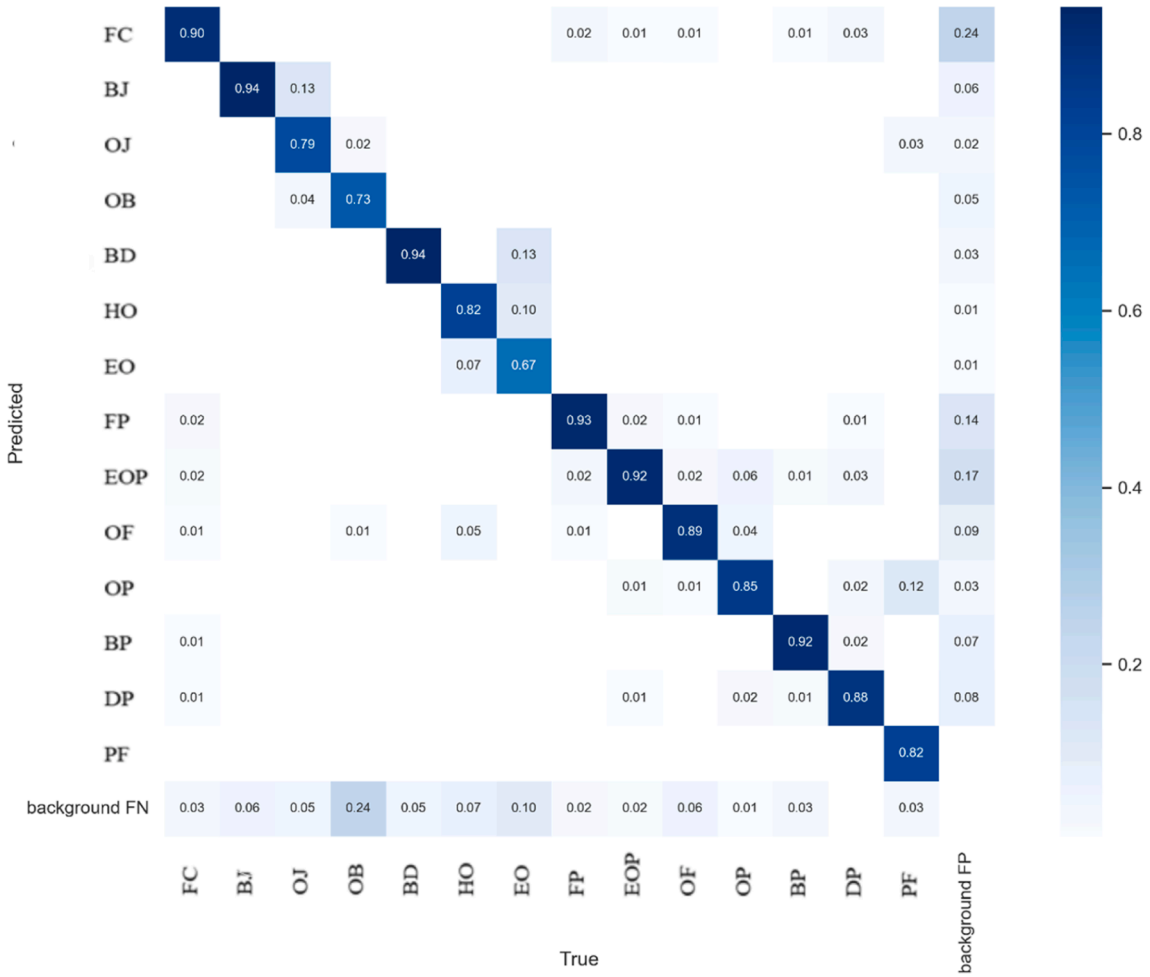


Fig. 8. Confusion matrix of 14 classes by YOLOv5l at a confidence threshold of 0.25. Column and row indicated predicted category and true category, respectively. The sum of values in each column equaled 1, while value in each row indicated proportion of predictions in corresponding category.

Table 2
Results from other studies on kiwifruit flower detection.

Author	Objects	Image pixel size	Model	AP of bud (%)	AP of flower (%)	mAP (%)	Model size (MB)	Average detection speed (ms/ image)
Williams et al. (2020)	Flower	1024 × 600	Faster R-CNN	—	—	85.30	142	100.00
Li et al. (2022)	Flower and bud	4608 × 3456	YOLOv4	98.57	96.66	97.61	244	38.64
Our study	Flower and bud	4608 × 3456	YOLOv5l	94.20	93.12	93.23	89.4	15.50

Note: For convenience of comparison, nine classes of multi-class flowers except bud were combined into one class (namely "Flower") in Table 2, and then mAP of nine classes was regarded as the AP of "Flower".

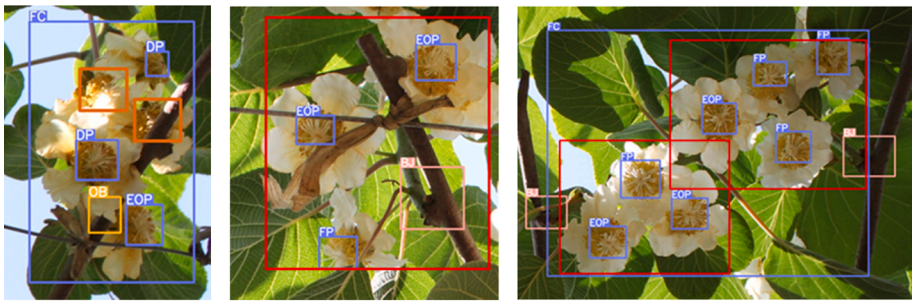
detection accuracy and detection speed with lower model size.

3.3. Matching results among multi-class objects

MA_t depends on consecutive twice matching, i.e., flower matching flower cluster and then flower cluster matching branch junction. For the consecutive twice matching, if one of them was wrong, the MA_t was affected a lot. MA_{fmc} was 97.60 %, which was 2.60 % higher than that of flower cluster matching branch junction. Although there was no direct connection between the value of MA_t and values of MA of twice matching, high values of the twice MA were the foundation of MA_t reaching a high value of 92.30 %. Overall, the MA_t met requirements of obtaining flower distributions.

Detection results of both multi-class flowers and flower clusters influenced MA_{fmc}. As shown in Fig. 9a, there were 5 flowers in a flower cluster, where 4 blue rectangles represented results of detected flowers matching detected flower cluster, but only three flowers and flower cluster were correctly detected. Besides, although three flowers were correctly detected in Fig. 9b, they were not detected as a flower cluster, where the true flower cluster was represented by a red rectangle. Maybe some flowers were too discrete to be detected as one flower cluster. As shown in Fig. 9c, two flower clusters were so close that they were detected as one flower cluster. However, situations of these two factors rarely happened in kiwifruit orchard with standardized management, which had little effect on results of multi-class flowers matching flower cluster.

As for MA_{cmj}, there were two main factors influencing it, i.e., detection results of flower clusters and that of branch junctions.



(a) (b) (c)

Fig. 9. Examples of missed and incorrect detection resulting in incorrect matching results. (a) Navy blue rectangles represented all correct detection results, while missed detection results were represented by orange rectangles; (b) Although three flowers were correctly detected, they were not detected as a true flower cluster, which was represented by a red rectangle; (c) Red rectangles represented true results of flower clusters, while a blue rectangle with "FC" label was incorrect, where only two true flower clusters were incorrectly detected one flower cluster. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Incorrect or missed detection might lead to no matched flower cluster or no matched branch junction. As shown in Fig. 10, a branch junction was almost completely occluded by a leave, and it could not be detected, which failed in flower cluster matching branch junction. Although correct matching was highly dependent on accurate detection, these situations also rarely happened and influenced MA_{cmj} little.

3.4. Matching results based on improved matching method

Although original match method achieved the clear majority of correct matching results, there was still an obvious matching error that needed to be solved in very small number of images. If two or many flower clusters matched the same branch junction based on the original matching method, there was only one flower cluster matching the branch junction, while the other flower cluster matched no branch junction, as shown in Fig. 11a. Thus, Euclidean distances of global matching results needed to be considered. If two or more flower clusters matched the same branch junction based on their minimum Euclidean distances, respectively, the flower cluster was selected to match the branch junction according to the minimum value of their minimum Euclidean distances. Examples of matching results based on improved matching method were shown in Fig. 11b.

All MAs increased except that of flowers matching flower clusters. MA_{fmc} based on improved matching method was same as that of original matching method due to improved matching method only on flower cluster matching branch junction. As shown in Table 3, MA_{cmj} (96.20 %) and MA_t (93.30 %) just increased by 1.20 % and 1.00 % based on improved matching method, respectively. That was because situation in Fig. 11 rarely existed in kiwifruit orchards under standardized management. Although both MA_{cmj} and MA_t increased just a little, there was no doubt that the improved matching method was effective to solve the match error of original matching method.

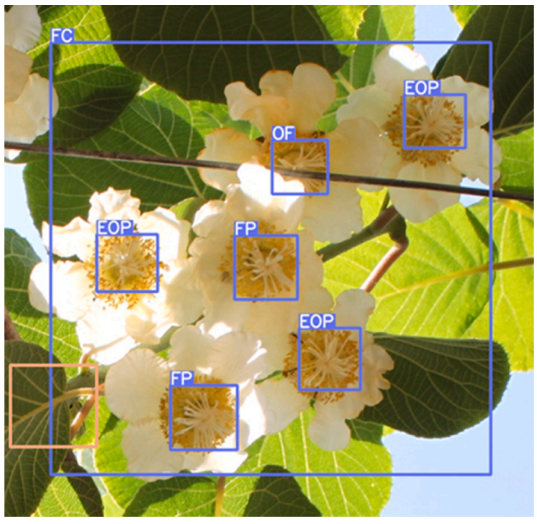


Fig. 10. Examples of incorrect matching results about flower cluster matching branch junction. All blue rectangles represented those correct objects that were detected by YOLOv5l, while a light pink rectangle represented a branch junction without being detected. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.5. Processing speed of detection and objects matching

YOLOv5 not only solved slow detection speed for a two-stage network such as Faster R-CNN but also had a more flexible network structure than that of previous YOLO models. As shown in Table 2, YOLOv5l attained satisfactory detection accuracy and faster detection speed (15.50 ms per image at 4068 × 3456 pixels) with lower model size.

In test results, total processing speed of multi-class flower detection and its distribution identification was 112.46 ms per image including 15.50 ms for image detection by YOLOv5l. According to a previous study by Duke et al. (2017) about robotic pollination, the possible minimum time for robotic pollinating was 500 ms per flower, including 50 ms for image processing. There was no doubt that the proposed method met these requirements.

4. Conclusions

Influences of flower phenology and flower distribution in fruit branch were not considered on pollination in most previous studies about robotic pollination of kiwifruit. However, indiscriminate pollination for almost all open flowers resulted in pollen wastage and not enough commodity kiwifruits. Thus, this work proposed a fast method of multi-class detection of kiwifruit flower and its distribution identification based on YOLOv5l and Euclidean distance, which was expected to choose suitable flowers for robotic pollination. Kiwifruit flowers were classified into 10 classes based on their phenology to find flowers in optimal pollination timing, while flower cluster and branch junctions were divided into 4 classes for obtaining flower distributions. After all classes were manually labeled by rectangular bounding boxes, YOLOv5l was used for training and detection. YOLOv5l attained satisfactory detection results, especially on 10 classes of kiwifruit flowers, which showed that more objects in training and objects classified into more classes might achieve better detection results. All Mas of both original matching method and improved matching method were above 92.00 %, which indicated that it's effective to obtain flower distributions using YOLOv5l and Euclidean distance under kiwifruit orchard with standardized management. Furthermore, total processing speed of the proposed method was 112.46 ms per image including 15.50 ms for image detection by YOLOv5l, which met requirements of processing and detection time for robotic pollination. Results showed that multi-class kiwifruit flowers and relative flower distributions could be accurately and fast obtained, which were helpful to gain suitable flowers for robotic pollination. Future improvements will be made by not only collecting

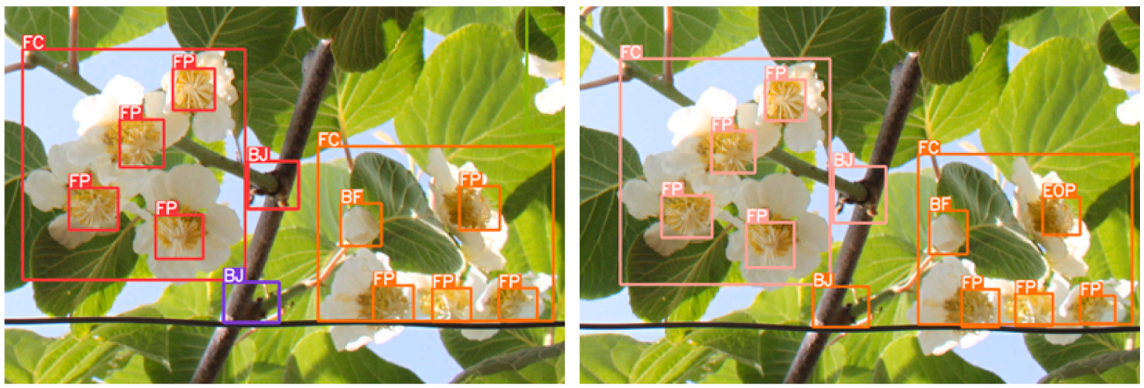


Fig. 11. Examples of matching results. (a) Original matching results, where red rectangles, orange rectangles, and a purple rectangle represented correct total matching results (among flowers, flower cluster and branch junction), correct results of only flowers matching flower cluster, and branch junction without matching, respectively; (b) Matching results based on improved matching method, where pink rectangles and orange rectangles represented their respective correct total matching results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Results from original method and improved matching method.

Method	MA_{fmc}	MA_{cmj}	MA_t
Original matching method	97.60 %	95.00 %	92.30 %
Improved matching method	97.60 %	96.20 %	93.30 %

images of kiwifruit flower in different species and locations to enlarge training dataset, but also further improving matching method based on real distance among multi-class flowers, flower clusters and branch junctions.

CRediT authorship contribution statement

Guo Li: Data curation, Investigation, Methodology, Writing – original draft. **Longsheng Fu:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Changqing Gao:** Conceptualization, Methodology, Writing – review & editing. **Wentai Fang:** Conceptualization, Methodology, Writing – review & editing. **Guanao Zhao:** Investigation, Conceptualization, Writing – review & editing. **Fuxi Shi:** Investigation, Methodology, Writing – review & editing. **Jaspreet Dhupia:** Methodology, Supervision, Writing – review & editing. **Kegang Zhao:** Conceptualization, Data curation, Methodology, Supervision, Writing – review & editing. **Rui Li:** Methodology, Writing – review & editing. **Yongjie Cui:** Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was partially supported by the Key R&D Program of Zhejiang Province, China (2022C02055); National Natural Science Foundation of China (32171897); Youth Science and Technology Nova Program in Shaanxi Province of China (2021KJXX-94); Science and Technology Promotion Program of Northwest A&F University (TGZX2021-29); China Postdoctoral Science Foundation funded project (2019M663832); National Foreign Expert Project, Ministry of Science and Technology, China (DL2022172003L, QN2022172006L).

References

Bochkovskiy, A., Wang, C., Liao, H.M., 2020. YOLOv4: Optimal speed and accuracy of object detection. arXiv Prepr. arXiv: 2004.10934.
Cangi, R., Atalay, D., 2018. Effects of different bud loading levels on the yield, leaf and fruit characteristics of Hayward kiwifruit. Hortic. Sci. 33, 23–28. https://doi.org/10.17221/3736-hortsci.
Castro, H., Siopa, C., Casais, V., Castro, M., Loureiro, J., Gaspar, H., Castro, S., 2021. Pollination as a key management tool in crop production: Kiwifruit orchards as a study case. Sci. Hortic. 290, 110533. https://doi.org/10.1016/j.scienta.2021.110533.
Dias, P.A., Tabb, A., Medeiros, H., 2018. Apple flower detection using deep convolutional networks. Comput. Ind. 99, 17–28. https://doi.org/10.1016/j.compind.2018.03.010.
Fu, L., Feng, Y., Wu, J., Liu, Z., Gao, F., Majeed, Y., Al-Mallahi, A., Zhang, Q., Li, R., Cui, Y., 2021. Fast and accurate detection of kiwifruit in orchard using improved YOLOv3-tiny model. Precis. Agric. 22 (3), 754–776. https://doi.org/10.1007/s11119-020-09754-y.
Gao, F., Fu, L., Zhang, X., Majeed, Y., Li, R., Karkee, M., Zhang, Q., 2020. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN. Comput. Electron. Agric. 176, 105634. https://doi.org/10.1016/j.compag.2020.105634.
Gianni, T., Vania, M., 2018. Artificial pollination in kiwifruit and olive trees. Pollinat. Plants. https://doi.org/10.5772/intechopen.74831.
Gonzalez, M.V., Coque, M., Herrero, M., 1995. Stigmatic receptivity limits the effective pollination period in kiwifruit. J. Am. Soc. Hortic. Sci. 120, 199–202. https://doi.org/10.21273/jashs.120.2.199.
Jocher, G., Stoken, A., Borovec, J., NanoCode012, ChristopherSTAN, Changyu, L., Laughing, Hogan, A., lorenzomammana, tkianai, yxNONG, AlexWang1900, Diaconu, L., Marc, wanghaoyang0106, ml5ah, Doug, Hatovix, Poznanski, J., L.Y., changyu98, Rai, P., Ferriday, R., Sullivan, T., Xinyu, W., YuriRibeiro, Claramunt, E.R., hopesala, pritul dave, yzchen, 2020. ultralytics/yolov5: v3.0. https://doi.org/10.5281/ZENODO.3983579.
Kuznetsova, A., Maleva, T., Soloviev, V., 2020. Using YOLOv3 algorithm with pre- and post-processing for apple detection in fruit-harvesting robot. Agronomy 10 (7), 1016. https://doi.org/10.3390/agronomy10071016.
Li, G., Suo, R., Zhao, G., Gao, C., Fu, L., Shi, F., Dhupia, J., Li, R., Cui, Y., 2022. Real-time detection of kiwifruit flower and bud simultaneously in orchard using YOLOv4 for robotic pollination. Comput. Electron. Agric. 193, 106641. https://doi.org/10.1016/j.compag.2021.106641.
Lim, J.Y., Ahn, H.S., Nejati, M., Bell, J., Williams, H., MacDonald, B.A., 2020. Deep neural network based real-time kiwi fruit flower detection in an orchard environment. arXiv Prepr. arXiv: 2006.04343.
Lin, F., Hou, T., Jin, Q., You, A., 2021. Improved YOLO based detection algorithm for floating. Entropy. 23 (9), 1111. https://doi.org/10.3390/e23091111.
Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature pyramid networks for object detection. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 936–944. https://doi.org/10.1109/CVPR.2017.106.
Liu, S., Qi, L., Qin, H., Shi, J., Jia, J., 2018. PANet: Path Aggregation Network for Instance Segmentation. In: 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake, pp. 8759–8768. doi: 10.1109/CVPR.2018.00913.
Majeed, Y., Karkee, M., Zhang, Q., 2020. Estimating the trajectories of vine cordons in full foliage canopies for automated green shoot thinning in vineyards. Comput. Electron. Agric. 176, 105671. https://doi.org/10.1016/j.compag.2020.105671.

McPherson, H.G., Richardson, A.C., Snelgar, W.P., Patterson, K.J., Currie, M.B., 2001. Flower quality and fruit size in kiwifruit (Actinidia deliciosa). New Zeal. J. Crop Hortic. Sci. 29, 93–101. https://doi.org/10.1080/01140671.2001.9514167.

Nepal, U., Eslamiat, H., 2022. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. Sensors 22 (2), 464.

Redmon, J., Farhadi, A., 2018. YOLOv3: An incremental improvement. arXiv Prepr. arXiv: 1804.02767.

Salinero, M.C., Vela, P., Sainz, M.J., 2009. Phenological growth stages of kiwifruit (Actinidia deliciosa 'Hayward'). Sci. Hortic. (Amsterdam) 121, 27–31. https://doi.org/10.1016/j.scienta.2009.01.013.

Smith, G.S., Gravett, I.M., Edwards, C.M., Curtis, J.P., Buwalda, J.G., 1994. Spatial analysis of the canopy of kiwifruit vines as it relates to the physical, chemical and postharvest attributes of the fruit. Ann. Bot. https://doi.org/10.1006/anbo.1994.1012.

Song, Z., Zhou, Z., Wang, W., Gao, F., Fu, L., Li, R., Cui, Y., 2021. Canopy segmentation and wire reconstruction for kiwifruit robotic harvesting. Comput. Electron. Agric. 181, 105933. https://doi.org/10.1016/j.compag.2020.105933.

Suo, R., Gao, F., Zhou, Z., Fu, L., Song, Z., Dhupia, J., Li, R., Cui, Y., 2021. Improved multi-classes kiwifruit detection in orchard to avoid collisions during robotic picking. Comput. Electron. Agric. 182, 106052. https://doi.org/10.1016/j.compag.2021.106052.

Thakur, A., Chandel, J.S., 2004. Effect of thinning on fruit yield, size and quality of kiwifruit cv. allison. Acta Hortic. 662, 359–364. https://doi.org/10.17660/actahortic.2004.662.53.

Tian, Y., Yang, G., Wang, Z., Li, E., Liang, Z., 2020. Instance segmentation of apple flowers using the improved mask R-CNN model. Biosyst. Eng. 193, 264–278. https://doi.org/10.1016/j.biosystemseng.2020.03.008.

Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M., 2020a. Scaled-YOLOv4: Scaling cross stage partial network. arXiv Prepr. arXiv: 2011.08036v1.

Wang, C.Y., Mark Liao, H.Y., Wu, Y.H., Chen, P.Y., Hsieh, J.W., Yeh, I.H., 2020b. CSPNet: A new backbone that can enhance learning capability of CNN. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. 2020-June, 1571–1580. https://doi.org/10.1109/CVPRW50498.2020.00203.

Wang, X., Tang, J., Whitty, M., 2021. DeepPhenology: Estimation of apple flower phenology distributions based on deep learning. Comput. Electron. Agric. 185, 106123. https://doi.org/10.1016/j.compag.2021.106123.

Williams, H., Nejati, M., Hussein, S., Penhall, N., Lim, J.Y., Jones, M.H., Bell, J., Ahn, H. S., Bradley, S., Schaare, P., Martinsen, P., Alomar, M., Patel, P., Seabright, M., Duke, M., Scarfe, A., MacDonald, B., 2020. Autonomous pollination of individual kiwifruit flowers: Toward a robotic kiwifruit pollinator. J. F. Robot. 37, 246–262. https://doi.org/10.1002/rob.21861.

Wu, Z., Li, G., Yang, R., Fu, L., Li, R., Wang, S., 2022. Coefficient of restitution of kiwifruit without external interference. J. Food Eng. 327, 111060. https://doi.org/10.1016/j.jfoodeng.2022.111060.

Wu, Z., Yang, R., Gao, F., Wang, W., Fu, L., Li, R., 2021. Segmentation of abnormal leaves of hydroponic lettuce based on DeepLabV3+ for robotic sorting. Comput. Electron. Agric. 190, 106443. https://doi.org/10.1016/j.compag.2021.106443.

Xu, R., Lin, H., Lu, K., Cao, L., Liu, Y., 2021. A forest fire detection system based on ensemble learning. Forests 12, 1–17. https://doi.org/10.3390/f12020217.

Zhang, J., Karkee, M., Zhang, Q., Zhang, X., Yaqoob, M., Fu, L., Wang, S., 2020. Multi-class object detection using Faster R-CNN and estimation of shaking locations for automated shake-and-catch apple harvesting. Comput. Electron. Agric. 173, 105384. https://doi.org/10.1016/j.compag.2020.105384.

Zhang, X., Karkee, M., Zhang, Q., Whiting, M.D., 2021. Computer vision-based tree trunk and branch identification and shaking points detection in Dense-Foliage canopy for automated harvesting of apples. J. F. Robot. 38, 476–493. https://doi.org/10.1002/rob.21998.