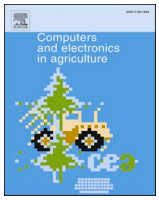
Computers and Electronics in Agriculture 229 (2025) 109715



Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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



Original papers



Growth characteristics based multi-class kiwifruit bud detection with overlap-partitioning algorithm for robotic thinning

Haojie Dang [a], Leilei He [a], Yufei Shi [a], Lamin L. Janneh [a,g], Xiaojuan Liu [a], Chi Chen [a], Rui Li [a], Hongbao Ye [b,*], Jinyong Chen [c], Yaqoob Majeed [h], Xiaoxi Kou [a,*], Longsheng Fu [a,d,e,f,*]

[a] College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
[b] Key Laboratory of Agricultural Equipment for Hilly and Mountainous Areas in Southeastern China (Co-construction by Ministry and Province), Ministry of Agriculture and Rural Affairs, Hangzhou, Zhejiang 310021, China
[c] Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou, Henan 450009, China
[d] Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China
[e] Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China
[f] Northwest A&F University Shenzhen Research Institute, Shenzhen, Guangdong 518000, China
[g] School of Information Communication and Technology, University of The Gambia, Banjul, Peace Building, Kanifing 3530, Gambia
[h] Department of Electrical Engineering and Computer Science, University of Wyoming, Laramie, WY 82071, United States

ARTICLE INFO

Keywords:
Main bud
Lateral bud
Deep learning
Overlap-partitioning algorithm
Different annotation strategies

ABSTRACT

Bud thinning is a critical operation in the early stage of kiwifruit production, which is currently performed by skilled workers and has an urgent need to develop bud thinning robots. Accurate detection of kiwifruit buds is the first step, which focuses on distinguishing main bud and lateral bud. Kiwifruit buds are small, with similar shapes and colors in the main and lateral buds. Therefore, two kiwifruit bud detection methodologies were proposed to distinguish them. One is two-stage kiwifruit bud detection methodology (T-SKBDM) with an enhanced algorithm that leverage kiwifruit bud growth characteristics after network training for precise detection of main and lateral buds, and another is one-stage kiwifruit bud detection methodology (O-SKBDM) that classifies buds during the training. These methodologies adopted a two-classes annotation strategy (T-CAS) and a five-classes annotation strategy (F-CAS), respectively. In addition, both utilized an overlap-partitioning algorithm (OPA) that partitions large images into small images with overlapping areas. YOLOv8l model was trained on the dataset with different annotation strategies before and after using the OPA. Results showed that the T-CAS achieved a mean average precision (mAP) of 66.4 % before employing the OPA, which was 17.5 % higher than the F-CAS. With the OPA, mAPs of the T-CAS and F-CAS increased by 15.8 % and 17.7 %, respectively. Furthermore, T-SKBDM improved by 12.0 % and 14.3 % in distinguishing main and lateral buds, respectively, compared with the average precisions of 69.2 % and 66.2 % for O-SKBDM. These results indicate that the T-SKBDM assists in detecting kiwifruit buds and distinguishing the main and lateral buds, thus laying the foundation for robotic bud thinning.

1. Introduction

Bud thinning is a critical early operation in kiwifruit production that aims to reduce nutrient competition and ensure fruit quality. The unique growth pattern of kiwifruit is characterized by a short flowering period and a long bud period, and this requires more attention to bud thinning than flower thinning (Zhou et al., 2023). Without bud thinning, there will be excess fruit growth with tremendously deteriorating quality (Pescie and Strik, 2004; Cangi and Atalay, 2006; Kumarihami et al., 2021). Current bud thinning practices require skilled workers to distinguish and manually remove the lateral bud from the canopy while retaining the main bud; these practices are cost-demanding, labor-intensive, and not sustainable for large-scale farming (Wouters et al., 2015; Mao et al., 2024; Xue et al., 2024). With the annual rise in labor costs and the aging population, there is an urgent need for automated bud thinning (Li et al., 2023; Tao et al., 2023). Consequently, the advancement of robotic bud thinning holds great promise in research.

Accurate detection of kiwifruit buds is essential for robotic bud

* Corresponding authors at: Institute of Agricultural Equipment, Zhejiang Academy of Agricultural Sciences, Hangzhou 310000, China (H. Ye); College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China (X. Kou, L. Fu).
E-mail addresses: yhb2008@zaas.ac.cn (H. Ye), kouxiaoxi@nwafu.edu.cn (X. Kou), fulsh@nwafu.edu.cn (L. Fu).

https://doi.org/10.1016/j.compag.2024.109715
Received 20 August 2024; Received in revised form 9 November 2024; Accepted 26 November 2024
Available online 6 December 2024
0168-1699/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

thinning. Due to the recent revolution of deep learning based on convolutional neural networks, researchers have applied deep learning networks to detect flowers and fruits in kiwifruit production. Fu et al. (2019) obtained an average precision (AP) of 92.3 % using Faster Region-Convolutional Neural Network (Faster R-CNN) with Zeller Fergus Net (ZFNet) for kiwifruit detection in an orchard. As Faster R-CNN does not meet the real-time requirement, Suo et al. (2021) used You Only Look Once (YOLO) series to detect multi-class kiwifruit, reaching a higher mean AP (mAP) of 91.9 % with YOLOv4. Li et al. (2022) utilized YOLO version five large (YOLOv5l) with transfer learning for multi-class detection of kiwifruit flower, which achieved mAP of 91.6 %. Similarly, Gao et al. (2023) employed YOLOv5l to detect kiwifruit flower, which reported a success detection rate (the number of flowers detected correctly divided by the number of all flowers) of 99.3 %. Despite these advancements in detecting flowers and fruits of kiwifruit, the detection of smaller objects, such as kiwifruit bud, remains a challenge (Liu et al., 2021; Mirzaei et al., 2023).

Current methods for small object detection mainly focus on modifying widely used deep learning network architectures. Wang et al. (2021) built a YOLO-Dense by adding a dense connection module to YOLOv3, which improved mAP by 8.1 % compared to the original YOLOv3 network for tomato anomalies detection in the field. Li et al. (2021) used a high-resolution backbone network in Faster-RCNN network and achieved mAP of 86.2 % for hydroponic lettuce seedlings, outperforming other detectors such as Fully Convolutional One-Stage (FCOS) and Single Shot MultiBox Detector (SSD). In the same vein, Tian et al. (2023) proposed a V-space based Multi-Scale Feature Fusion SSD network for apple leaf disease detection, which improved mAP by 8.7 % compared to the original SSD network. These studies focus on modifying network for small object detection of crops, while few studies pay attention to data enhancement for improving detection accuracy (Zhou et al., 2019; Tong et al., 2020). Despite the reported promising results of these methods, network modifications are different for each specific object, which leading to a lack of replicability and thus requiring a universal method.

Overlap-partitioning strategy has proven efficient and effective in detecting small objects within complex backgrounds and large fields of view. It is a strategy based on data enhancement that partitions large images into small images. Bao et al. (2023) cropped a wheat RGB image acquired by an Unmanned Aerial Vehicle (UAV) from both positive and negative directions to highlight wheat lesions, thereby detecting wheat scab and achieving an AP of 80.4 %. Jiang et al. (2023) proposed image overlap-partitioning and stitching for BlendMask to segment wire images acquired in a modern apple orchard, which was 38.4 % higher than the results obtained from a full image dataset. Liu et al. (2024) used a similar approach to detect kiwifruit and its calyx in images that collected from different cameras, which improved mAP by 6.2 %. Due to the high detection accuracy of above researches for small objects, the strategy was embraced for detecting kiwifruit buds.

After detecting the kiwifruit buds, it is important to distinguish the main and lateral buds of the kiwifruit for robotic bud thinning. This specification enables the thinning operation to primarily focus on removing the lateral bud while preserving the main bud (Thompson, 2014; Romano et al., 2019). In addition, the main and lateral buds typically grow in bud clusters (McNeilage, 1991; Watson and Gould, 1993), and within the bud cluster, they exhibit similar colors and shapes, whereas the size of lateral bud is somewhat smaller than main bud, which is a crucial distinguishing factor between main bud and lateral bud. Therefore, this study will distinguish between main and lateral buds based on their growth characteristics.

In this study, a growth characteristics based multi-class kiwifruit bud detection methodology with overlap-partitioning algorithm (OPA) for robotic thinning was proposed. It integrates the OPA with the cutting-edge YOLO model to precisely detect kiwifruit buds. Furthermore, the methodology capitalizes on the accurate detection of kiwifruit buds to locate the target edge box and compare its size, thereby distinguishing between main and lateral buds. This study paves the way for more precise and efficient kiwifruit thinning practices by addressing the challenges of bud detection and differentiation.

2. Materials and methods

Two different detection methodologies are proposed in this study to accurately detect kiwifruit buds and distinguish between main and lateral buds. The first one is a two-stage kiwifruit bud detection methodology (T-SKBDM) based on bud growth characteristics as shown in the red rectangular box of Fig. 1. The other is a one-stage kiwifruit bud detection methodology (O-SKBDM), as shown in the green rectangular box of Fig. 1. These methodologies both employed OPA to improve the detection accuracy. For a more comprehensive understanding, a detailed schematic diagram illustrating the workflow of these methodologies is provided in Fig. 1.

2.1. Image acquisition

Images of 'Xuxiang' kiwifruit buds were acquired during March to April of 2022 and 2023 from International Kiwifruit Innovation Orchard (34°17′ N, 108°2′ E, approximately 504 m in altitude) in Yangling, Shaanxi province, China. This kiwifruit variety is widely planted in this area, and most of its buds hang down from the canopy towards the ground. The acquisition of images requires capturing the proper view of the buds not only for detection but also for mimicking the robot field operation view. Thus, a camera is placed parallel to the canopy and captured images from bottom to top. Considering to the cost burden of agricultural robots as well as the operational needs of field conditions, an ordinary RGB-D camera (RealSense D435, Intel) was chosen to collect images of buds at around 50 cm below the canopy with resolutions of 1,920 × 1,080 pixels. A total of 245 original images were acquired at different times in the morning and afternoon under different lighting conditions, which were saved in Portable Network Graphics (PNG) format. Examples of images under different lighting conditions are shown in Fig. 2.

2.2. Image annotation and data preprocessing

Based on the growth characteristics of kiwifruit buds and the operational needs of bud thinning, kiwifruit buds were classified with a two-classes annotation strategy (T-CAS) or a five classes annotation strategy (F-CAS), respectively.

In the F-CAS, it is worth reiterating that bud thinning focuses on removing the lateral bud while preserving the main bud. The main and lateral buds grow in bud clusters, which are divided into double buds, triple buds and beyond. Therefore, kiwifruit buds were subdivided into five classes. The first class indicated a kiwifruit bud cluster that consists of three or more buds (referred to as three_buds in this study), as shown in Fig. 3(a1). The second class showed a kiwifruit bud cluster with two buds (referred to as two_buds), as shown in Fig. 3(a2). The third class, depicted in Fig. 3(b1), represented the main bud of the kiwifruit bud cluster (referred to as main_bud). The fourth class, illustrated in Fig. 3 (b2), denoted the lateral bud in the kiwifruit bud cluster (referred to as lateral_bud). The last class indicated single kiwifruit buds (referred to as single_bud), as shown in Fig. 3(b3).

In the T-CAS, because kiwifruit buds grow in clusters with similar shape and color, the buds were divided into two classes. The first class contains Fig. 3(a1) and Fig. 3(a2), which indicate kiwifruit buds are clustered (referred to as bud_cl in this study), as shown in Fig. 3(a). The second class contains Fig. 3(b1), Fig. 3(b2) and Fig. 3(b3) indicating each kiwifruit bud (referred to as bud), which is shown in Fig. 3(b).

Kiwifruit buds were manually labeled with boxes that were tangent to kiwifruit bud outlines, as shown in Fig. 4. Acquired images were manually annotated by the graphical image annotation tool LabelImg (https://github.com/tzutalin/labelImg) that generated corresponding

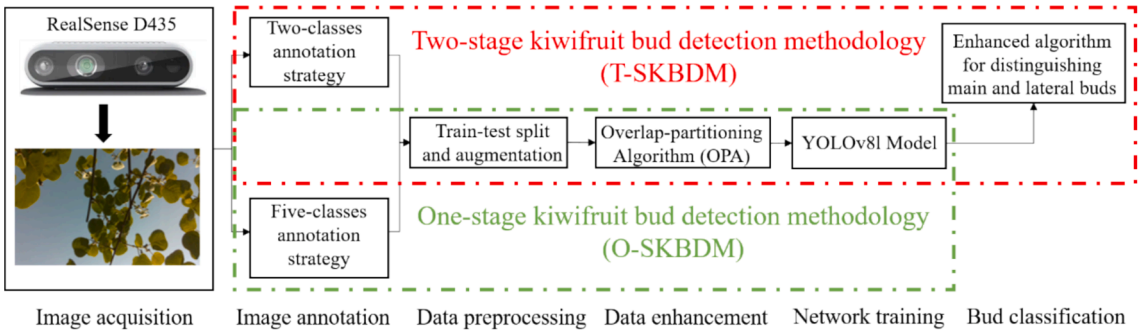


Fig. 1. Overall schematic workflow. T-SKBDM is in the red rectangular box, and O-SKBDM is in the green rectangular box.



Fig. 2. Examples of kiwifruit bud images acquired in orchard with (a) weak-lighting condition, (b) strong-lighting condition, and (c) back-lighting condition.

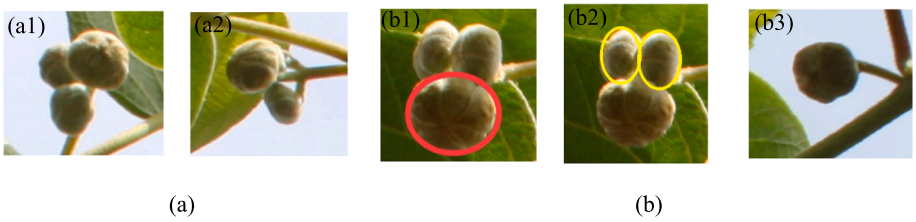


Fig. 3. Multi-classes kiwifruit bud images, F-CAS containing (a1), (a2), (b1), (b2), and (b3) and T-CAS containing (a) snd (b). (a1) Three or more kiwifruit buds (three_buds); (a2) Two kiwifruit buds (two_buds); (b1) Main bud of kiwifruit bud cluster in red circle (main_bud); (b2) Lateral bud of kiwifruit bud cluster in yellow circles (lateral_bud); (b3) Single kiwifruit bud (single_bud); (a) Kiwifruit bud cluster (bud_cl); (b) Kiwifruit bud (bud).



Fig. 4. Examples of kiwifruit buds labeled into (a) T-CAS, where buds inside purple and light green boxes referred to the bud_cl and bud, respectively; and (b) F-CAS, where buds in green, yellow, red, blue and orange boxes represented the three_buds, two_buds, main_bud, lateral_bud and single_bud, respectively.

label files including kiwifruit bud classes and box coordinates. For the T-CAS, kiwifruit buds inside purple and light green boxes represented the bud_cl and bud, respectively, as shown in Fig. 4(a). For the F-CAS, kiwifruit buds in green, yellow, red, blue and orange boxes represented the three_buds, two_buds, main_bud, lateral_bud and single_bud, respectively, as shown in Fig. 4(b). XML annotation files were generated after labeling, including folder name, image name, image path, image size, class name, and pixel coordinates of the label boxes.

After being annotated, the image dataset was split and augmented for training. This image dataset (245 images) was randomly divided into a train set (171 images), a validation set (49 images) and a test set (25 images) with a 7:2:1 ratio. Then, ten dataset augmentation methods were adopted to improve overall model performance, including brightness transformation, contrast transformation, chroma transformation, sharpness transformation, image mirroring in horizontal and vertical axes, and image rotation in 90° and 180°, where the transformation thresholds for brightness, contrast, and chroma were set to 0.7 and 1.3, respectively. After dataset augmentation, the training set was augmented from 171 images to 1,881 images.

2.3. Data enhancement

Kiwifruit buds occupy a small proportion of pixels in the original image, which poses a challenge to current object detection models based deep learning network. For this reason, an OPA is proposed to enhance data and increase the proportion of bud pixels in the images. The principle of the algorithm is to partition the original large image from top-left to bottom-right into small partitioned images of the training network input pixels. Some partitioned images may not have kiwifruit buds while the others increase the proportion of bud pixels. Generally, partitioning the images into a series of pixel patches will truncate certain parts of the target at the edge of adjacent partitions, resulting in the loss of some critical visual information and impacting the network training.

Therefore, an overlap rate is set in adjacent partitioned images to prevent missing kiwifruit bud information at the edge of the partitions. As shown in Fig. 5, three black rectangular boxes are overlapped to indicate that there is an overlap rate among three adjacent partitioned images, where the red area represents the overlapping regions.

The relative coordinates of the images in the edge region during image partitioning will exceed parts of the original image, resulting in a white space in certain areas of the partitioned image. In order to avoid this situation, the relative coordinates of the edge region need to be shifted. The diagram of the partitioned image shifting is shown in Fig. 6. The blue box indicates the original calculated partitioning position of the image, and the red box indicates the corrected partitioning position of the image. In this paper, the partitioned small image is set to 640 × 640 pixels for adapting with network input size.

XML annotation files must be partitioned simultaneously using the OPA to avoid exhaustive patch annotation after image partitioning. This involves training and detecting small images and corresponding annotation files after each partition and merging the detection results of small images after partitioning into large images.

2.4. Network training

You Only Look Once version 8 large (YOLOv8l) demonstrates significant potential in object detection within machine vision. It outperforms the other three counterparts within the YOLOv8 series in terms of accuracy (Sportelli et al., 2023). The enhanced precision of this model makes it suitable for tasks requiring feature detail analysis, such as identifying small and dense objects like kiwifruit buds (Terven et al., 2023). In light of its superior performance metrics, this research has specifically selected YOLOv8l to determine the presence of kiwifruit buds. In retrospect, YOLOv8l magnificent detection ability will effectively and efficiently strengthen the findings of this study.

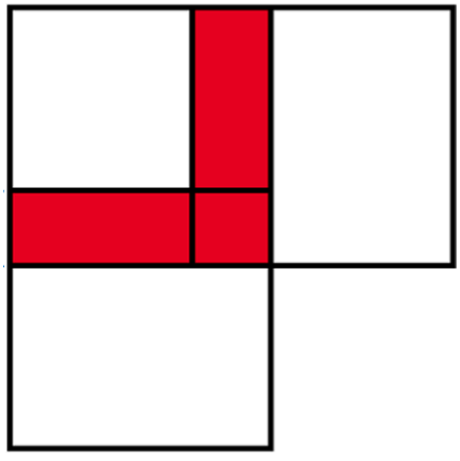


Fig. 5. Example of the overlap rate setting and overlapping regions is in the red area.

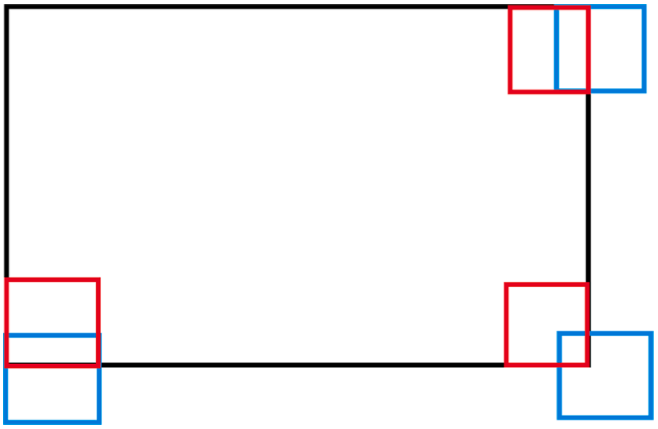


Fig. 6. Example of partitioned image shifting.

Experiments were performed based on PyTorch framework with version 2.0.1 based on a desktop computer equipped with Intel Xeon Gold 5128 32-core processor (2.3 GHz), dual Nvidia GeForce GTX 4090 24G GPUs, and 128 GB of memory, running on a 64-bit Windows 10 system. The software toolkit includes CUDA 12.1, cuDNN 8.8.1, Python 3.9, and OpenCV 4.8. YOLOv8l was applied to train kiwifruit bud detection on the PyTorch framework. The network input size was 640 × 640 pixels, with a batch size of 16. Stochastic gradient descent was used for training, with a momentum of 0.937 and a weight decay of 0.0005. A value of 0.01 was set as initial learning rate of the network, and epochs were set to 300 for analyzing training process.

2.5. Bud classification

In the T-CAS of kiwifruit buds for this study, no initial distinction was made between the main and lateral buds. It is imperative to realize that the primary objective of bud thinning in kiwifruit cultivation is to selectively remove the lateral buds while preserving the more crucial main buds. Despite their similar color and shape, which makes differentiation challenging in a cluster, the main buds typically exhibit a larger size than their lateral counterparts. Based on this observation, our methodology incorporated a bud classification enhanced algorithm to categorize the buds within a cluster accurately. This algorithm utilized the actual differences in size and the spatial positioning of the buds relative to the part of the cluster as tangible criteria to effectively distinguish between the main and lateral buds.

Firstly, the model predictive information from the image frame is translated into an XML file. The XML file contains annotations that delineate bud and bud_cl, each defined by bounding box coordinates (x_{min}, y_{min}, x_{max}, y_{max}). Then, the centroid coordinates (C_x, C_y) of each bud were calculated using Eqs. (1) and (2). This centroid calculation facilitates the correct determination of whether a bud resides within a given bud_cl by comparing its centroid with the bounding box of the cluster.

$$C_x = \frac{x_{min} + x_{max}}{2} \tag{1}$$

$$C_y = \frac{y_{min} + y_{max}}{2} \tag{2}$$

Finally, multi-class kiwifruit buds in the same bud_cl were compared in size according to the bounding box area (*BBA*) calculation shown in Eq. (3), and the largest was labeled as the main bud, while the rest were labeled as lateral buds. The results of the enhanced algorithm for bud classification are shown in Fig. 7, where the two-classes of buds before distinguishing and the four-classes of buds after distinguishing are shown in Fig. 7(a) and 7(b), respectively. This enhanced algorithm ensures that each bud_cl is accurately annotated with its constituent buds

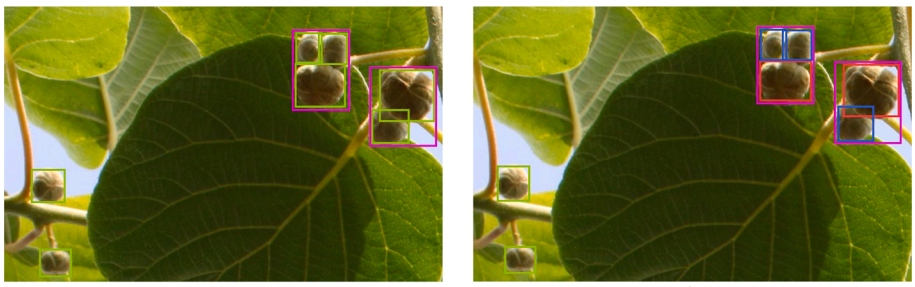


Fig. 7. The results of the enhanced algorithm for bud classification. (a) Before distinguishing bud (where buds inside purple and light green boxes referred to the bud_cl and bud, respectively); (b) After distinguishing bud (where buds in light green, purple, red and blue boxes represented the bud, bud_cl, main_bud and lateral_bud, respectively).

categorized appropriately, aligning with the practicality of kiwifruit bud thinning.

$$BBA = (x_{max} - x_{min}) \times (y_{max} - y_{min}) \tag{3}$$

2.6. Performance evaluation

The performance evaluation was conducted using the average precision (AP_i) for each class, the mean average precision (mAP_k), and the average detection speed. Specifically, AP_i for each class was derived using precision (P_i) and recall (R_i), as detailed in Eqs. (4) and (5), respectively. In the context of the F-CAS, the variable i corresponds to each specific class: three_buds ($i = 1$), two_buds ($i = 2$), main_bud ($i = 3$), lateral_bud ($i = 4$), and single_bud ($i = 5$). Similarly, in the T-CAS and four-classes scenarios, the meaning of i parallels that of the five-classes setup.

$$P_i = \frac{TP_i}{TP_i + FP_i} \tag{4}$$

$$R_i = \frac{TP_i}{TP_i + FN_i} \tag{5}$$

In this study, True Positives (TP_i) denote the quantity of correctly identified kiwifruit buds in the i^{th} class, while False Positives (FP_i) correspond to the count of erroneously detected kiwifruit buds in the i^{th} class. False Negatives (FN_i), on the other hand, indicate the number of kiwifruit buds that the model failed to detect in the i^{th} class.

As defined in Eq. (6), AP_i is the area under the curve of the P_i and R_i. AP measures the sensitivity of YOLOv8l model in detecting targets and also acts an index that reflects the performance of the network. Defined in Eq. (7), mAP_k represents the mean AP value across the k classes of kiwifruit buds. This metric comprehensively evaluates the performance of YOLOv8l model across multiple classes.

$$AP_i = \int_0^1 P_i(R_i)\mathrm{d}R_i \tag{6}$$

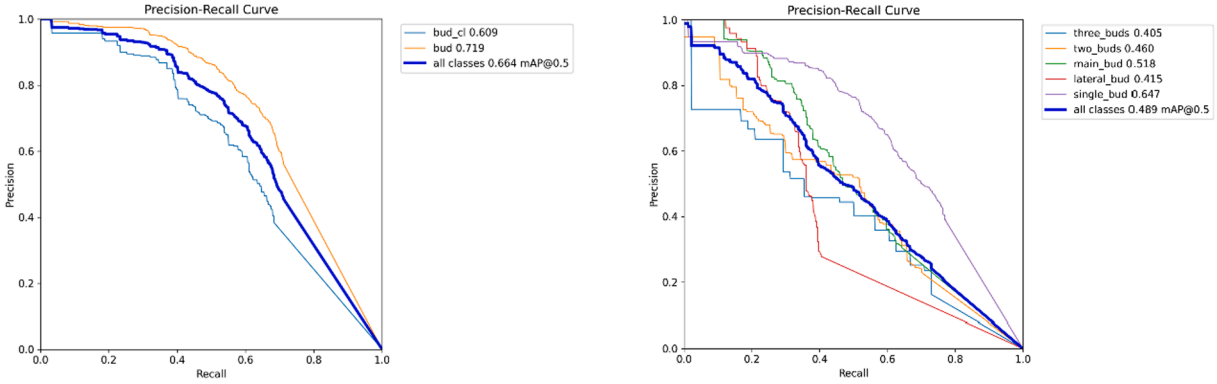
$$mAP_k = \frac{1}{k}\sum_{i=1}^{k} AP_i \tag{7}$$

3. Results and discussion

3.1. Detection comparison of the F-CAS and T-CAS

YOLOv8l model was combined with different annotation strategies to detect kiwifruit buds, which showed that T-CAS was superior to F-CAS. Multi-classes kiwifruit bud detection results of YOLOv8l model in the T-CAS and F-CAS are shown in Fig. 8(a) and (b), respectively. The F-CAS resulted in APs of 40.5 %, 46.0 %, 51.8 %, 41.5 %, and 64.7 % for for each sub-class, leading to an overall mAP of 48.9 %. Similarly, each class in the T-CAS had APs of 60.9 % and 71.9 % in the sub-class, while improving the mAP by 17.5 %. Furthermore, the performance of trained models was also evaluated on test set, which presented mAPs of 48.9 % and 67.0 % by F-CAS and T-CAS, respectively, approximated to the training results (as shown in Table 1). The comparison indicates that the simpler T-CAS not only simplifies labeling task but also provides more reliable and precise results.

The T-CAS is better than the F-CAS, which is closely related to the growth characteristics of kiwifruit buds. These are relatively small and share a similar color with the surrounding foliage, further complicating the detection process (Mirzaei et al., 2023). When network is tasked with distinguishing between finely detailed classes like main_bud and lateral_bud, its ability to generalize effectively is reduced due to fewer accessible characteristics. This visual similarity causes a higher degree



(a) (b)

Fig. 8. Result of multi-classes kiwifruit bud detection in the (a) T-CAS and (b) F-CAS by YOLOv8l.

Table 1
Result of test set kiwifruit bud detection in the F-CAS and T-CAS.

Annotation strategy	*AP* (%)					*mAP* (%)
F-CAS	three_buds	two_buds	main_bud	lateral_bud	single_bud	48.9
	39.8	47.2	51.2	40.3	66.2	
T-CAS	bud_cl		bud			67.0
	60.5		73.4			

of misclassification within the F-CAS, ultimately lowering its overall detection accuracy. The main_bud and lateral_bud classes have subtle distinctions in subtle biological distinction, yet machine learning models struggle to leverage that subtle difference effectively. The similarities in appearance among main_bud, lateral_bud and single_bud led to confusion during the training and inference phases.

In this study, an interesting finding is that reducing annotation classes for small objects like kiwifruit buds with complex backgrounds and visual similarity can improve detection performance and further reduce complexity. However, Suo et al. (2021) classified fruits into four and five categories based on robotic picking strategy and field occlusions, which showed that mAP of fruits in the five-classes is higher than that in the four-classes, and reported that increasing the labeling and training classes of the fruit could improve mAP. These claims cannot be generalized, especially when the objects of interest are small and exhibit only subtle variations. Thus, opting for a less detailed classification scheme may be more beneficial. In summary, it is important to make different annotation strategies depending on the object itself and its environment.

3.2. Benefit of the adopted OPA in kiwifruit bud detection

The adopted OPA has proven to significantly enhance the accuracy of YOLOv8l network in detecting and classifying kiwifruit buds. The results obtained before adopting the OPA align with those described in the previous section (Section 3.1), as shown in Fig. 8. The result of both annotation strategies in terms of AP and mAP was not encouraging. After adopting the OPA, multi-classes kiwifruit bud detection results of YOLOv8l model in the T-CAS and F-CAS are shown in Fig. 9(a) and (b), respectively. Compared with the results before adopting the OPA, APs increased by 18.4 %, 14.0 %, 17.4 %, 24.7 % and 14.2 % in the F-CAS. Likewise, APs were increased by 16.2 % and 15.4 % in the T-CAS. The mAP for F-CAS improved from 48.9 % to 66.6 % with an increasement of 17.7 %. Similarly, the T-CAS was improved by 15.8 % in mAP. Performance of enhanced models by the OPA on the test set was presented in Table 2. The mAPs achieved 66.1 % and 84.2 % for F-CAS and T-CAS, respectively, which improved by 17.2 % and 16.8 % compared to without the OPA. Overall, the stability and reliability of models were significantly improved after adopting the OPA.

Another notable aspect is that implementing the OPA results in slower detection speeds. The average detection speeds of the T-CAS and F-CAS were 88.5 ms and 98.2 ms, respectively, about 80 ms slower than without the OPA. The primary reason for this is the additional computational overhead introduced by partitioning large images into small images with overlapping areas. However, in practical agricultural environments, this trade-off is acceptable. For instance, fruit picking robots utilize a two-step strategy involving the detection of fruits in the camera field of view, followed by path planning for the robotic arm and end-effector to pick fruits (Gao et al., 2020). It has been reported that a kiwifruit picking robot equipped with a single-arm gripper requires approximately 2.78 s to pick one kiwifruit (Williams et al., 2020).

Robotic bud thinning employs a similar two-step strategy that involves detecting buds in the camera field of view, followed by path planning for the robotic arm and end-effector to thin buds. Due to the special trellis cultivation mode of kiwifruit, the operation paths of the robotics of fruit picking and bud thinning are similar with a bottom-up motion trajectory. Therefore, it is acceptable to spend 98.2 ms to detect a kiwifruit bud image and improve the detection accuracy in variant illumination and complex orchard environments. To learn more features for network, the large image is partitioned into smaller images that match the size of the network input, thereby improving overall detection reliability. Although it introduces some computational complexity, its high detection accuracy in real-world environments makes it a valuable algorithm.

3.3. Buds distinction in O-SKBDM and T-SKBDM

When comparing the O-SKBDM and T-SKBDM for detecting main and lateral kiwifruit buds, it is evident that the latter is superior. The O-SKBDM involves annotating the buds into five-classes and then using YOLOv8l model with the OPA for detection. It resulted in APs of 69.2 % and 66.2 % for main and lateral buds, respectively, while mAP was 66.6 %, as shown in Fig. 9(b). In contrast, the T-SKBDM annotated the buds into two-classes and utilizes YOLOv8l model with the OPA. It subsequently distinguished between main and lateral buds according to the enhanced algorithm based on the growth characteristics of kiwifruit

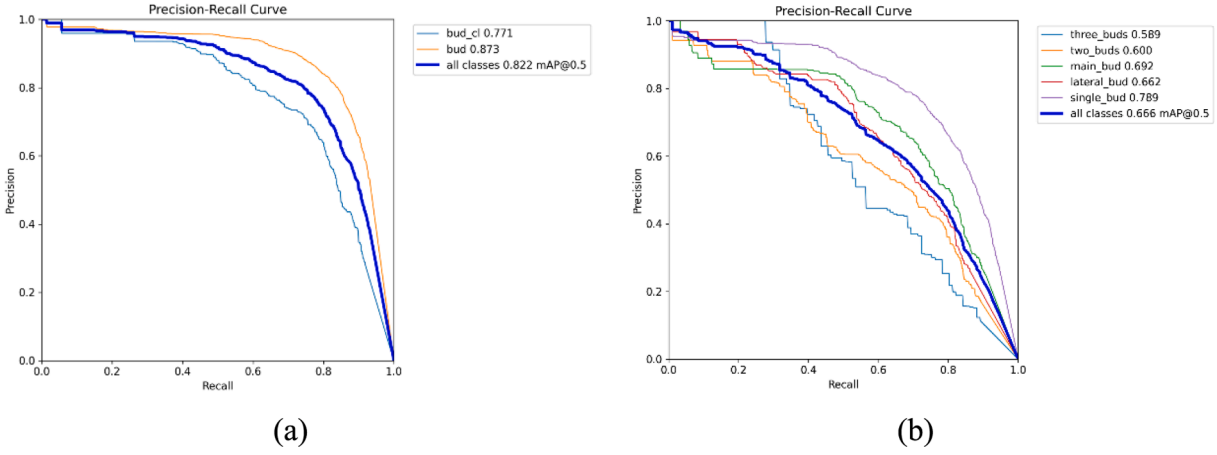


Fig. 9. Results of multi-classes kiwifruit bud detection in the (a) T-CAS and (b) F-CAS by YOLOv8l after adopting OPA.

Table 2
Result of test set kiwifruit bud detection in the F-CAS and T-CAS after adopting OPA.

Annotation strategy	*AP* (%)					*mAP* (%)
F-CAS	three_buds	two_buds	main_bud	lateral_bud	single_bud	66.1
	58.7	59.2	68.4	66.6	77.8	
T-CAS	bud_cl		bud			83.8
	78.9		88.6			

buds. The APs of the T-SKBDM were calculated to be 81.2 % and 80.5 % for main and lateral buds, with mAP reaching 81.4 %. The average time to process each file during the enhanced algorithm operation is 22 ms. This additional processing time is necessary to achieve the improved detection accuracy in distinguishing between main and lateral buds. This methodology not only reduces complexity of the preliminary classification labeling but also refines the subsequent identification process, leading to higher detection accuracy.

To visualize differences in the detection of main and lateral buds between the two methodologies, results were tested on same image in the test set, as shown in Fig. 10(a) and (b). Errors in detection are marked manually with black rectangles for false positives and white rectangles for false negatives. According to Fig. 10, the T-SKBDM resulted in two false positives, while the O-SKBDM had one false positive and three false negatives. The primary reason for these false negatives is the visual similarity between kiwifruit buds and the surrounding canopy, which complicates the detection process. False positives occur because kiwifruit buds are small objects, and the main and lateral buds are similar colors and shapes, differing mostly in size. This minimal difference is difficult for the network to learn, leading to misclassification. Taken together, the T-SKBDM is still superior to the O-SKBDM on kiwifruit bud detection.

Based on the growth characteristics of kiwifruit buds, the enhanced algorithm improved the reliability and accuracy of distinguishing main and lateral buds. The O-SKBDM tries to distinguish main and lateral buds from the beginning, resulting in poor accuracy. On the other hand, the T-SKBDM first accurately detects the buds and then distinguishes between main and lateral buds by applying the enhanced algorithm based on their growth characteristics. This methodology ensures that the model focuses on learning the finer details of the buds, reducing the chances of misclassification and improving overall detection. Moreno et al. (2023) proposed a new method to detect weeds using YOLOv8l and RetinaNet, with mAP reaching 93 % under the most challenging conditions where weed species may overlap. However, none of these models can accurately detect weeds when they are crowded together at a particular point. This scenario is similar to the distinction between the main and lateral bud, which may be solved if similar T-SKBDM is applied. This morphologically similar and dense scenario of weeds is similar to distinguishing main and lateral bud scenarios, where direct multi-classification detection does not work well to distinguish objects. In this study, O-SKBDM attempted to detect main and lateral buds by direct multi-classification, which led to unsatisfactory results. T-SKBDM first simplified the classification task, and then utilized the growth characteristics under vertical imaging to distinguish between main and lateral buds, which led to higher detection accuracy.

4. Conclusions

This study utilizes YOLOv8l model, combining different annotation strategies and algorithms to detect and classify kiwifruit buds. A comparison of the T-CAS and the F-CAS before and after using the OPA showed that mAP of the T-CAS before using the OPA was 66.4 %, which was 17.5 % higher than the F-CAS. After using the OPA, mAPs of T-CAS and F-CAS increased to 82.2 % and 66.6 %, respectively. According to the result, it may be more beneficial to choose an annotation strategy with fewer categories when the objects are small and have only minor differences. The OPA improved detection performance for both annotation strategies, proving its effectiveness for detecting small objects in images within complex background and large field of view. Compared to the O-SKBDM, the findings also highlight the superiority of the T-SKBDM with an enhanced algorithm based on growth characteristics in distinguishing main and lateral buds. In addition, accurate detection of main and lateral buds provides the possibility of the subsequent implementation of intelligent kiwifruit bud thinning equipment. However, the best mAP achieved for the kiwifruit bud detection model was only 82.2 %, indicating room for improvement. To enhance detection accuracy, expanding the dataset by collecting data across multiple growing years could be an effective approach, improving the robustness



Fig. 10. Examples of kiwifruit bud detection by (a) T-SKBDM and (b) O-SKBDM. False positives were marked by black rectangles manually drawn, while false negatives by white rectangles.

and generalization of model to various factors in field conditions.

CRediT authorship contribution statement

Haojie Dang: Writing – original draft, Methodology, Investigation, Data curation. **Leilei He:** Writing – review & editing, Software, Investigation. **Yufei Shi:** Writing – review & editing, Software. **Lamin L. Janneh:** Writing – review & editing, Data curation. **Xiaojuan Liu:** Writing – review & editing, Methodology. **Chi Chen:** Writing – review & editing, Methodology, Investigation. **Rui Li:** Writing – review & editing, Investigation. **Hongbao Ye:** Writing – review & editing, Methodology, Conceptualization. **Jinyong Chen:** Writing – review & editing, Methodology. **Yaqoob Majeed:** . **Xiaoxi Kou:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization. **Longsheng Fu:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization.

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.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (32371999); Science and Technology Program of Xi'an City, China (24NYGG0035); Open Project of Key Laboratory of Agricultural Equipment for Hilly and Mountainous Areas in Southeastern China (Co-construction by Ministry and Province), Ministry of Agriculture and Rural Affairs, China (QSKF2023002); Key Research and Development Program in Henan Province, China (221111111800); National Foreign Expert Project, Ministry of Science and Technology, China (QN2022172006L, DL2022172003L).

Data availability

Data will be made available on request.

References

Bao, W., Liu, W., Yang, X., Hu, G., Zhang, D., Zhou, X., 2023. Adaptively spatial feature fusion network: an improved UAV detection method for wheat scab. Precis. Agric. 24, 1154–1180. https://doi.org/10.1007/s11119-023-10004-0.

Cangi, R., Atalay, D.A., 2006. 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.

Fu, L., Tola, E., Al-Mallahi, A., Li, R., Cui, Y., 2019. A novel image processing algorithm to separate linearly clustered kiwifruits. Biosyst. Eng. 183, 184–195. https://doi.org/10.1016/j.biosystemseng.2019.04.024.

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.

Gao, C., He, L., Fang, W., Wu, Z., Jiang, H., Li, R., Fu, L., 2023. A novel pollination robot for kiwifruit flower based on preferential flowers selection and precisely target. Comput. Electron. Agric. 207, 107762. https://doi.org/10.1016/j.compag.2023.107762.

Jiang, H., Sun, X., Fang, W., Fu, L., Li, R., Cheein, F.A., Majeed, Y., 2023. Thin wire segmentation and reconstruction based on a novel image overlap-partitioning and stitching algorithm in apple fruiting wall architecture for robotic picking. Comput. Electron. Agric. 209, 107840. https://doi.org/10.1016/j.compag.2023.107840.

Kumarihami, H.M.P.C., Park, H.G., Kim, S.M., Park, J.I., Lee, E.J., Kim, H.L., Kim, J.G., 2021. Flower and leaf bud density manipulation affects fruit set, leaf-to-fruit ratio, and yield in southern highbush 'Misty' blueberry. Sci. Hortic. 290, 110530. https://doi.org/10.1016/j.scienta.2021.110530.

Li, G., Fu, L., Gao, C., Fang, W., Zhao, G., Shi, F., Dhupia, J., Zhao, K., Li, R., Cui, Y., 2022. Multi-class detection of kiwifruit flower and its distribution identification in orchard based on YOLOv5l and Euclidean distance. Comput. Electron. Agric. 201, 107342. https://doi.org/10.1016/j.compag.2022.107342.

Li, K., Gong, W., Shi, Y., Li, L., He, Z., Ding, X., Wang, Y., Ma, L., Hao, W., Yang, Z., Cui, Y., 2023. Predicting positions and orientations of individual kiwifruit flowers and clusters in natural environments. Comput. Electron. Agric. 211, 108039. https://doi.org/10.1016/j.compag.2023.108039.

Li, Z., Li, Y., Yang, Y., Guo, R., Yang, J., Yue, J., Wang, Y., 2021. A high-precision detection method of hydroponic lettuce seedlings status based on improved Faster RCNN. Comput. Electron. Agric. 182, 106054. https://doi.org/10.1016/j.compag.2021.106054.

Liu, X., Jing, X., Jiang, H., Younas, S., Wei, R., Dang, H., Wu, Z., Fountas, S., Fu, L., 2024. Performance evaluation of newly released cameras for fruit detection and localization in complex kiwifruit orchard environments. J. f. Robot. 41 (4), 881–894. https://doi.org/10.1002/rob.22297.

Liu, Y., Sun, P., Wergeles, N., Shang, Y., 2021. A survey and performance evaluation of deep learning methods for small object detection. Expert Syst. Appl. 172, 114602. https://doi.org/10.1016/j.eswa.2021.114602.

Mao, W., Murengami, B., Jiang, H., Li, R., He, L., Fu, L., 2024. UAV-based high-throughput phenotyping to segment individual apple tree row based on geometrical features of poles and colored point cloud. J. ASABE 67 (5), 1231–1240. https://doi.org/10.13031/ja.15895.

McNeilage, M.A., 1991. Gender variation in Actinidia deliciosa, the kiwifruit. Sex. Plant Reprod. 4, 267–273. https://doi.org/10.1093/aob/mcr217.

Mirzaei, B., Nezamabadi-pour, H., Raoof, A., Derakhshani, R., 2023. Small object detection and tracking: A comprehensive review. Sensors 23, 6887. https://doi.org/10.3390/s23156887.

Moreno, H., Gómez, A., Altares-López, S., Ribeiro, A., Andújar, D., 2023. Analysis of stable diffusion-derived fake weeds performance for training convolutional neural networks. Comput. Electron. Agric. 214, 108324. https://doi.org/10.1016/j.compag.2023.108324.

Pescie, M.A., Strik, B.C., 2004. Thinning before bloom affects fruit size and yield of hardy kiwifruit. HortScience 39, 1243–1245. https://doi.org/10.21273/hortsci.39.6.1243.

Romano, A., Torregrosa, A., Balasch, S., Ortiz, C., 2019. Laboratory device to assess the effect of mechanical thinning of flower buds, flowers and fruitlets related to fruitlet developing stage. Agronomy 9, 5–14. https://doi.org/10.3390/agronomy9110668.

Sportelli, M., Apolo-Apolo, O.E., Fontanelli, M., Frasconi, C., Raffaelli, M., Peruzzi, A., Perez-Ruiz, M., 2023. Evaluation of YOLO object detectors for weed detection in different turfgrass scenarios. Appl. Sci. 13, 8502. https://doi.org/10.3390/app13148502.

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.

Tao, K., Wang, Z., Yuan, J., Liu, X., 2023. Design of a novel end-effector for robotic bud thinning of Agaricus bisporus mushrooms. Comput. Electron. Agric. 210, 107880. https://doi.org/10.1016/j.compag.2023.107880.

Terven, J., Córdova-Esparza, D.M., Romero-González, J.A., 2023. A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Mach. Learn. Knowl. Extr. 5, 1680–1716. https://doi.org/10.3390/make5040083.

Thompson, A.B., 2014. Determining the effective pollination period and effects of crop load reduction on Au kiwifruit cultivars. Auburn University.

Tian, L., Zhang, H., Liu, B., Zhang, J., Duan, N., Yuan, A., Huo, Y., 2023. VMF-SSD: A novel V-space based multi-scale feature fusion SSD for apple leaf disease detection. IEEE/ACM Trans. Comput. Biol. Bioinforma. 20, 2016–2028. https://doi.org/10.1109/TCBB.2022.3229114.

Tong, K., Wu, Y., Zhou, F., 2020. Recent advances in small object detection based on deep learning: A review. Image vis. Comput. 97, 103910. https://doi.org/10.1016/j.imavis.2020.103910.

Wang, X., Liu, J., 2021. Tomato anomalies detection in greenhouse scenarios based on YOLO-Dense. Front. Plant Sci. 12, 1–11. https://doi.org/10.3389/fpls.2021.634103.

Watson, M., Gould, K.S., 1993. The development of fruit shape in kiwifruit: Growth characteristics and positional differences. J. Hortic. Sci. 2, 185–194. https://doi.org/10.1080/00221589.1993.11516342.

Williams, H., Ting, C., Nejati, M., Jones, M.H., Penhall, N., Lim, J.Y., Seabright, M., Bell, J., Ahn, H.S., Scarfe, A., Duke, M., MacDonald, B., 2020. Improvements to and large-scale evaluation of a robotic kiwifruit harvester. J. f. Robot. 37, 187–201. https://doi.org/10.1002/rob.21890.

Wouters, N., De Ketelaere, B., Deckers, T., De Baerdemaeker, J., Saeys, W., 2015. Multispectral detection of floral buds for automated thinning of pear. Comput. Electron. Agric. 113, 93–103. https://doi.org/10.1016/j.compag.2015.01.015.

Xue, P., Li, Q., Fu, G., 2024. Design and control simulation analysis of tender tea bud picking manipulator. Appl. Sci. 14, 928. https://doi.org/10.3390/app14020928.

Zhou, H., Ou, J., Meng, P., Tong, J., Ye, H., Li, Z., 2023. Reasearch on kiwi fruit flower recognition for efficient pollination based on an improved YOLOv5 algorithm. Horticulturae 9, 400. https://doi.org/10.3390/horticulturae9030400.

Zhou, J., Vong, C.M., Liu, Q., Wang, Z., 2019. Scale adaptive image cropping for UAV object detection. Neurocomputing 366, 305–313. https://doi.org/10.1016/j.neucom.2019.07.073.