

An inter-row navigation line detection method for *Camellia oleifera* orchards based on UAVs and DSConv-YOLOv11

Linhui WANG^{1,2,3}, Xiaowen LIU³, Lizhi CHEN (✉)^{1,3}, Yonghong TAN (✉)^{1,3}, Bin QING^{3,4}, Yu YANG^{1,3}, Mengyu ZENG^{1,2,3}

1 School of Intelligent Manufacturing, Hunan University of Science and Engineering, Yongzhou 425199, China.

2 College of Agricultural Unmanned Systems, China Agricultural University, Beijing 100193, China.

3 Hunan Provincial Engineering Research Center for Information Perception and Early Warning of Smart Agriculture (Fruits and Vegetables), Yongzhou 425199, China.

4 Qidian Zongheng (Yongzhou) Technology Co., Ltd., Yongzhou 425199, China.

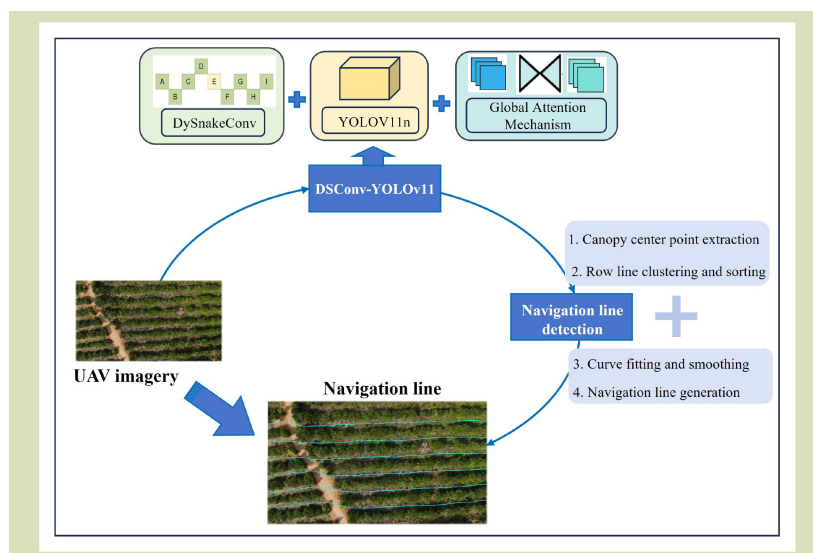
KEYWORDS

UAVs, YOLOv11, Navigation line, *Camellia oleifera*

HIGHLIGHTS

- An improved DSConv-YOLOv11 model is proposed, which effectively enhances the accuracy and local anti-interference stability of canopy detection under the complex partial interference conditions of actual *Camellia oleifera* orchard scenarios.
- A hybrid navigation line generation strategy combining K-Means clustering, IQR outlier removal and cubic spline fitting is designed to solve crop row discontinuity and curved row fitting problems.
- Multi-dimensional evaluation metrics including CRDA, heading deviation and lateral deviation are adopted to systematically analyze the correlation between canopy detection precision and navigation path quality.

GRAPHICAL ABSTRACT



ABSTRACT

Camellia oleifera orchard navigation presents considerable challenges, including the limited field of view of ground robots, discontinuous crop rows, and complex shadow interference. To address these challenges, this study proposes a UAV-based method for inter-row navigation line detection using an improved DSConv-YOLOv11 model. The Global Attention Mechanism is incorporated into the YOLOv11n backbone to enhance the model's ability to focus on canopy features, while the original C3k2 module is replaced with C3k2_DySnakeConv to improve the extraction of geometric features associated with curved canopies and inter-row gaps. Navigation lines are generated through K-Means clustering of canopy center points, outlier

Received July 2, 2026;
Accepted August 26, 2026.

Correspondences: clz0329@huse.edu.cn,
tyh2977@huse.edu.cn

Special Issue: Engineering Applications and
Practices of Agricultural UAVs in
Smart Agriculture

removal, and cubic spline fitting. The improved model achieves a mean average precision (mAP) of 85.3%, representing a 5.2-percentage-point improvement over YOLOv11n, with an inference time of 303.33 ms per aerial image. The generated navigation lines achieve an average crop row detection accuracy (CRDA) exceeding 0.73 compared with manually annotated paths, with a mean heading deviation of 5.25° and a lateral deviation of 11.62 pixels. Comparative experiments using a *C. oleifera* picking robot further demonstrate that the extracted navigation paths are generally consistent with the robot's actual driving trajectory and meet the requirements of inter-row operations. The proposed method provides a feasible technical solution for detecting inter-row navigation lines in *C. oleifera* orchards from a low-altitude remote-sensing perspective.

© The Author(s) 2027. Published by Higher Education Press. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0>)

1 Introduction

Agricultural mechanization and precision agriculture have become increasingly important for addressing global food security challenges and enhancing agricultural production efficiency^[1]. With continued global population growth and growing pressure on cultivated land resources, traditional extensive farming practices that rely heavily on human experience and manual labor are increasingly unable to meet the comprehensive requirements of modern agriculture for precision, sustainability, and economic viability. Autonomous navigation is a fundamental capability of agricultural robots, and its core function is to generate reliable navigation paths that accurately reflect actual field conditions while satisfying vehicle kinematic constraints^[2]. The accuracy and stability of these paths not only directly influence operational performance but also determine the feasibility and robustness of subsequent path-tracking control algorithms^[3]. Therefore, obtaining good and reliable paths under different field conditions, season and environmental conditions represents an essential scientific challenge in the field of intelligent agriculture.

At present, autonomous navigation paths for agricultural machinery are primarily obtained using three approaches. Global navigation satellite systems (GNSS) typically rely on operation trajectories that are pre-planned by operators according to field boundaries and geometric characteristics^[4]. Although this approach performs well in flat and relatively regular large-scale fields, it has limited adaptability when crop rows become curved or misaligned because of seeding errors, terrain undulations, or variations in crop growth. In such

conditions, GNSS-based navigation cannot directly perceive the actual spatial distribution of crops and therefore lacks the ability to adaptively adjust navigation paths online^[5]. LiDAR can provide high-precision three-dimensional information regarding terrain and obstacles, making it particularly promising for less structured environments, including orchards and protected agricultural facilities^[6,7]. However, the relatively high cost of LiDAR sensors and their susceptibility to interference under adverse weather conditions, such as rain and fog, restrict their large-scale deployment in typical agricultural environments, particularly in the hilly and mountainous regions of southern China^[8]. In contrast, vision-based navigation offers advantages including lower sensor costs, rich information acquisition, and considerable algorithmic scalability. Consequently, it has emerged as a major research focus and an important direction for autonomous navigation in agricultural machinery^[9–15].

Extensive research has been conducted on vision-based navigation path extraction, with existing approaches generally classified into object detection-based and semantic segmentation-based methods. For object detection-based approaches, Diao et al.^[16] developed an improved YOLOv8s network incorporating an Atrous Spatial Pyramid Pooling (ASPPF) structure to enhance multi-scale feature extraction. The proposed method accurately detected the central points of maize plants and fitted crop rows using the centre coordinates of the detection bounding boxes, achieving an average precision of 90.2%. Yin et al.^[17] developed a lightweight Maize-YOLO model for complex maize field environments and combined it with an IPM-DBSCAN-LSM framework to fit

crop-row centre lines. Their method achieved a positional accuracy of 98.6% and an angular deviation of 0.44°. Yang et al.^[18] employed a YOLO-based end-to-end network to directly predict traversable regions for agricultural machinery. The Excess Green operator and Otsu's adaptive thresholding method were subsequently applied to extract crop-row navigation lines, with a processing time of approximately 25 ms per frame, indicating its potential for real-time agricultural operations. Zheng et al.^[19] proposed YOLOv11-IMP, a robust multi-scale anchor-free multi-task detection method with further improvements to spatial pyramid pooling, large-kernel attention, and cross-modal vision-language feature fusion. For grape cluster detection, the precision and recall reached 94.3% and 93.5%, respectively, and the mean absolute error was 0.46 kg per grape vine.

Semantic segmentation-based methods have also attracted extensive research interest for navigation path extraction. Zhang et al.^[20] constructed a ResC-UNet model adopting ResNet-50 as the backbone and introducing an attention mechanism to strengthen the perception of slender targets and boundary regions, achieving a crop row segmentation accuracy of 94.9%. Li et al.^[21] adopted a Transformer-based semantic segmentation model taking rice stems as the main detection objects. The method effectively mitigates interference from irregular leaf growth and obtains a detection accuracy of 92.93%. Ying et al.^[22] employed YOLOv8-Seg to generate masks of strawberry ridges, and then adopted the Canny edge detection algorithm to extract contour lines. By statistically analyzing the intercept distribution of edge points traversed by slope lines with slopes of ± 1 , the peak positions of ridges were determined. These key points were subsequently connected to generate navigation lines, with an accuracy of 96%. Numerous other researchers have designed lightweight semantic segmentation networks to balance segmentation accuracy and inference speed when deployed on resource-limited embedded systems^[23,24].

Despite the favorable detection performance reported for the aforementioned methods in specific scenarios, several common limitations remain to be addressed. First, most existing approaches assume that crop rows are continuous and intact. Under actual field conditions, however, local plant loss, extensive weed growth, canopy occlusion, and substantial variations in illumination can introduce fractures, gaps, or indistinct boundaries into segmentation masks. These defects may result in abrupt local deviations in the fitted navigation

lines, thereby compromising the smoothness and reliability of navigation^[25]. Second, most existing studies have focused on approximately straight crop rows and have relied on linear functions or polyline approximations based on local endpoint connections for path fitting. Such approaches are inadequate for accurately representing the nonlinear and continuously curved geometry of crop rows^[26]. In orchards on hilly ground, terraced fields and locations with a natural undulating character, the datasets generated from the crops display considerable bending in crop rows causing large fitting errors in those sections of the curves and making it impossible for the standard algorithms to meet the needs of high-precision navigation control^[27].

To address the inherent limitations of single-view vision-based approaches, recent studies have increasingly incorporated low-altitude remote sensing technologies based on unmanned aerial vehicles (UAVs). UAVs can capture the global spatial distribution of crops from a broader bird's-eye perspective, thereby providing more comprehensive information for navigation path planning^[27-29]. Lin et al.^[28] combined UAV aerial imagery with a multi-task YOLO algorithm to jointly infer traversable-region segmentation and weed distribution, thereby improving the robustness of navigation-line extraction and achieving a coordinate conversion error of only 5.47 cm. Li et al.^[29] constructed three-dimensional oblique-photography models of agricultural fields using low-altitude remote sensing and integrated them with semantic segmentation to generate global three-dimensional navigation paths covering both operational areas and headland turning regions. This approach provides a new strategy for global path planning in complex terrain. In addition, the integration of UAVs with visual simultaneous localization and mapping (SLAM) technology has gradually been applied to navigation map construction in unstructured environments, including orchards^[30,31]. Collectively, these studies demonstrate that combining UAV remote sensing with deep learning can provide more robust solutions for navigation path construction in complex agricultural environments, particularly in economic forest and fruit-growing systems characterized by irregular planting patterns and substantial terrain undulations.

C. oleifera is an important woody oil crop in China and is predominantly cultivated in the low-hilly regions of southern China. Fragmented orchard terrain, irregular plant distribution, and complex canopy structures present substantial challenges to the autonomous navigation of

ground-based agricultural machinery. Existing ground-based vision navigation methods are particularly vulnerable to restricted fields of view, severe occlusion, and highly curved crop rows in such environments, making it difficult to achieve stable and accurate navigation-line detection under complex local orchard conditions. Moreover, research specifically addressing navigation path detection in *C. oleifera* orchards remains limited, and systematic, targeted technical solutions are still lacking.

Against this background, this study considers the global spatial distribution of *C. oleifera* planting areas from a bird's-eye perspective and accounts for the nonlinear geometric characteristics of inter-row crop distribution. Accordingly, an inter-row navigation-line detection method is proposed based on a UAV platform and an improved DSConv-YOLOv11 network. The main contributions of this study are as follows:

- (1) An improved DSConv-YOLOv11 model is proposed to enhance canopy detection accuracy and local anti-interference robustness under complex partial occlusion and interference conditions in actual *C. oleifera* orchards.
- (2) A navigation-line generation method integrating K-Means clustering, interquartile range (IQR)-based outlier removal, and cubic spline fitting is developed to address crop-row discontinuity and nonlinear curvature, thereby generating continuous and smooth inter-row navigation paths.
- (3) The influence of canopy detection accuracy on navigation-line quality is systematically evaluated using CRDA, heading deviation, and lateral deviation as key performance metrics.

2 Materials and methods

2.1 Image acquisition

The data used in this study were collected through the DJI Mini 5 Pro UAV (DJI Innovation, Shenzhen, China) in *C. oleifera* orchard planting base located in Yongzhou City, Hunan Province, China (longitude: 111°58' E, latitude: 26°18' N). The flights were performed on a sunny, cloudless weather with stable illumination in order to assure consistent high-quality images collected from the UAV at an altitude of 80 m with a resolution of 3840 × 2160 pixels. In total, 414 images were collected, and these were randomly split into training,

validation, and test sets in proportions of 8:1:1. The dataset included several difficult field conditions, including shadows, weeds interferences, inter-canopy occlusion, crop-row discontinuities due to roadways. Enlarged images show the camellia features in Fig. 1.

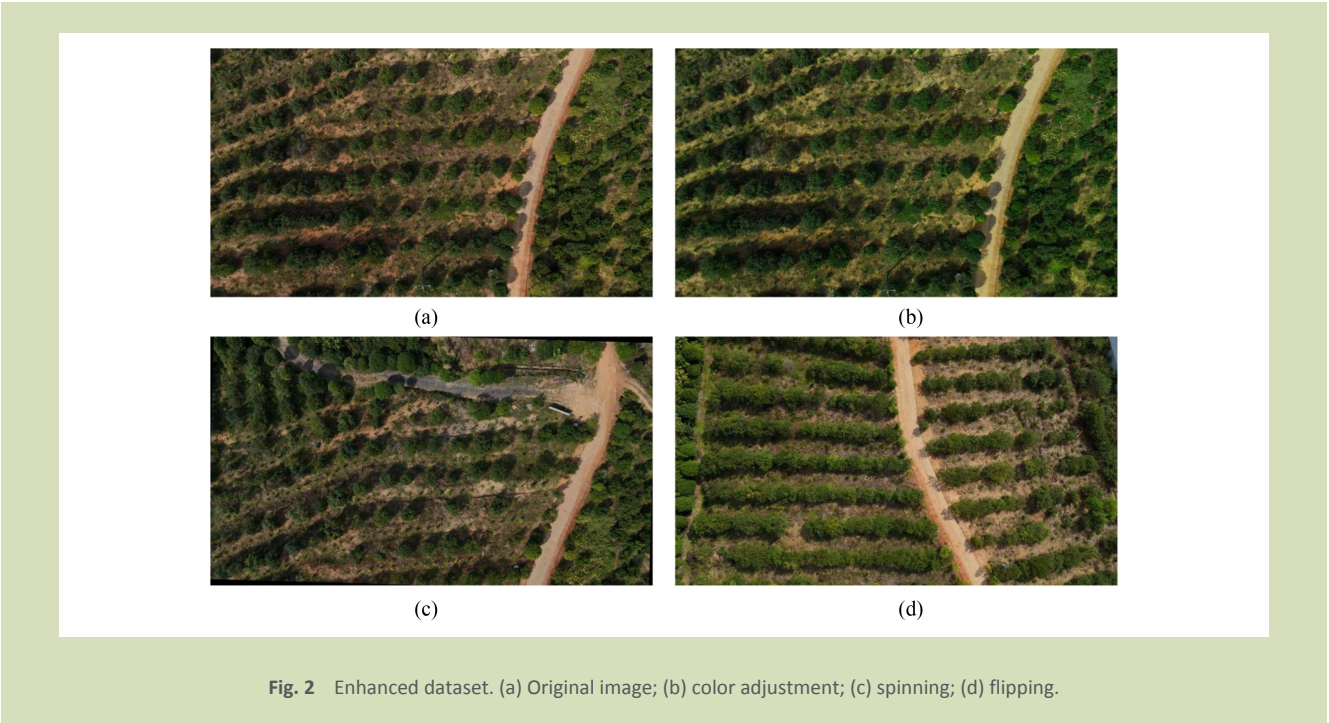
2.2 Dataset construction

At the start of model training, raw images were proportionally resized to 640 × 360 pixels to fit the input specifications of the YOLO model and shorten model training time. Images were resized using the resize function in the OpenCV library with Lanczos interpolation. The canopies of *C. oleifera* trees were manually annotated using the Make Sense annotation tool. Unified criteria were developed to guarantee accurate and consistent canopy identification. For canopies of individual *C. oleifera* trees, the outermost continuous canopy outline was marked as the annotation boundary, whereas small twigs and sparsely distributed leaves near canopy edges were excluded. Where canopies overlapped, adjoined one another, or were partially obscured, dividing boundaries were defined according to visible inter-canopy gaps and trunk distances to reduce ambiguity between neighboring canopies. All images were annotated independently by two researchers. The researchers compared annotations, resolved discrepancies and reached consensus to reduce subjectivity and ensure the accuracy and reliability of the dataset. The annotation data were first saved in XML format and then converted into TXT format. To improve dataset variability and enhance model generalizability, several data augmentation methods were applied during model training, including flipping, spinning and color adjustments. Ultimately, the augmented dataset contained a total of 1654 images, with representative examples shown in Fig. 2.

2.3 Navigation line detection method

2.3.1 Improved YOLOv11 model

In this research, we select YOLOv11n as the baseline model and implement targeted modifications to its feature extraction module and attention strategy. The modified model is named DSConv-YOLOv11, as shown in Fig. 3. To tackle the limited receptive field of the conventional C3k2 module, which struggles to accurately capture the curved contours of *C. oleifera* trees and inter-row gaps, we replace the original C3k2 module with DySnakeConv. The improved module



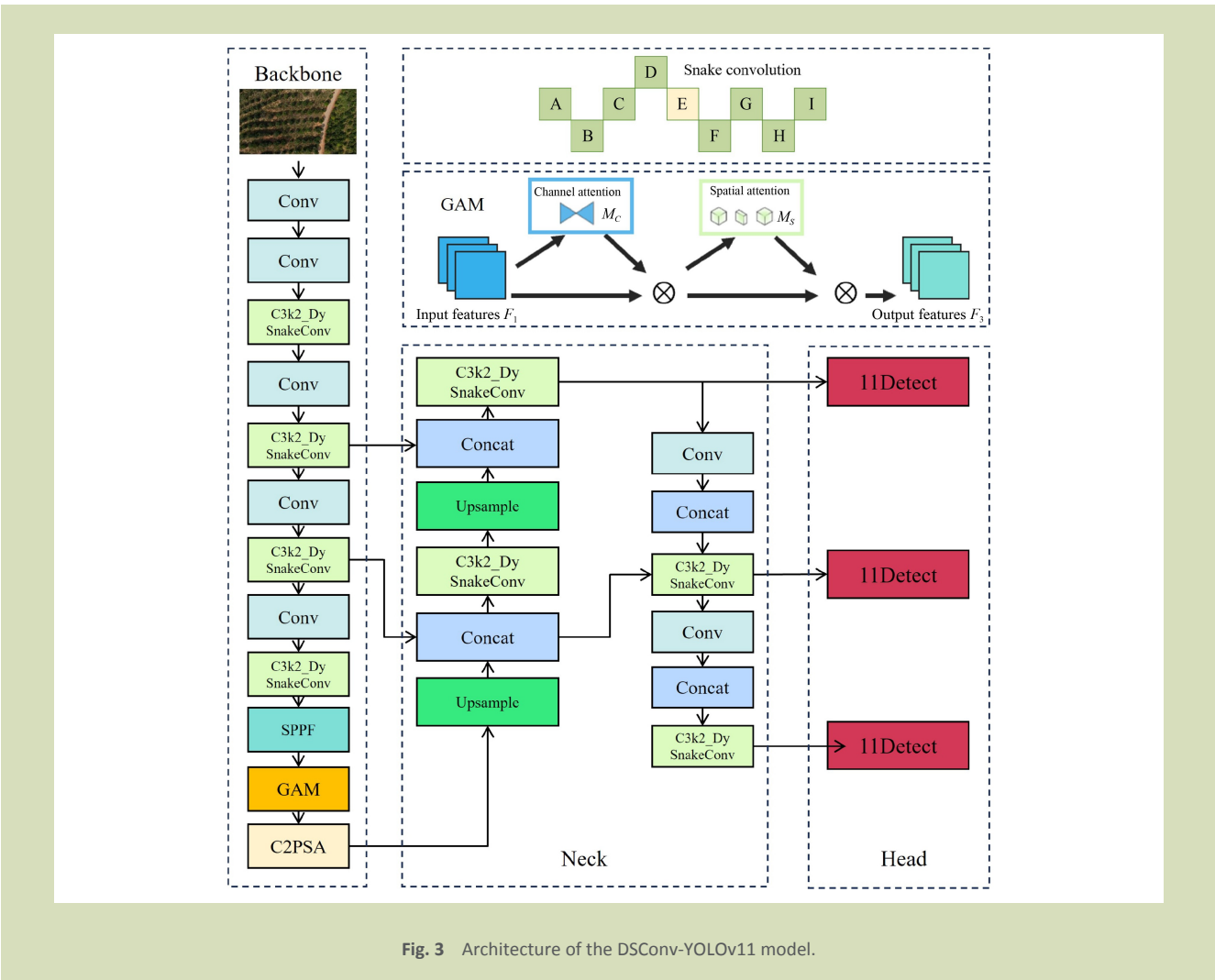


Fig. 3 Architecture of the DSConv-YOLOv11 model.

adopts multi-branch convolution and dynamic serpentine sampling to enhance the extraction of fine-grained features such as tree canopy outlines and inter-row intervals. This alleviates feature extraction uncertainties arising from complex terrain and canopy occlusion. Furthermore, we integrate the Global Attention Mechanism (GAM) into the network, enabling the backbone to retain global channel and spatial information and strengthen cross-dimensional feature interactions.

2.3.1.1 Attention mechanism improvement

To reduce background interference from weeds, shadows, and illumination variations in *C. oleifera* orchard scenes, the GAM was incorporated into the backbone network of YOLOv11^[32]. The structure of the GAM is illustrated in Fig. 4.

GAM employs a serial architecture comprising channel attention followed by spatial attention. Within the channel attention module, three-dimensional feature interactions and a multilayer perceptron are employed to adaptively enhance channels associated with *C. oleifera* canopy features while suppressing irrelevant information from background elements such as weeds and soil. The spatial attention module incorporates a 7×7 convolution to model the spatial distribution and morphological characteristics of the canopy more effectively, thereby preserving contour details that may otherwise be lost through pooling operations. Compared with conventional attention mechanisms, GAM improves the model's ability to capture discriminative features under complex conditions, including inter-canopy occlusion and variations in terrain and canopy scale in hilly orchard environments, while maintaining spatial information.

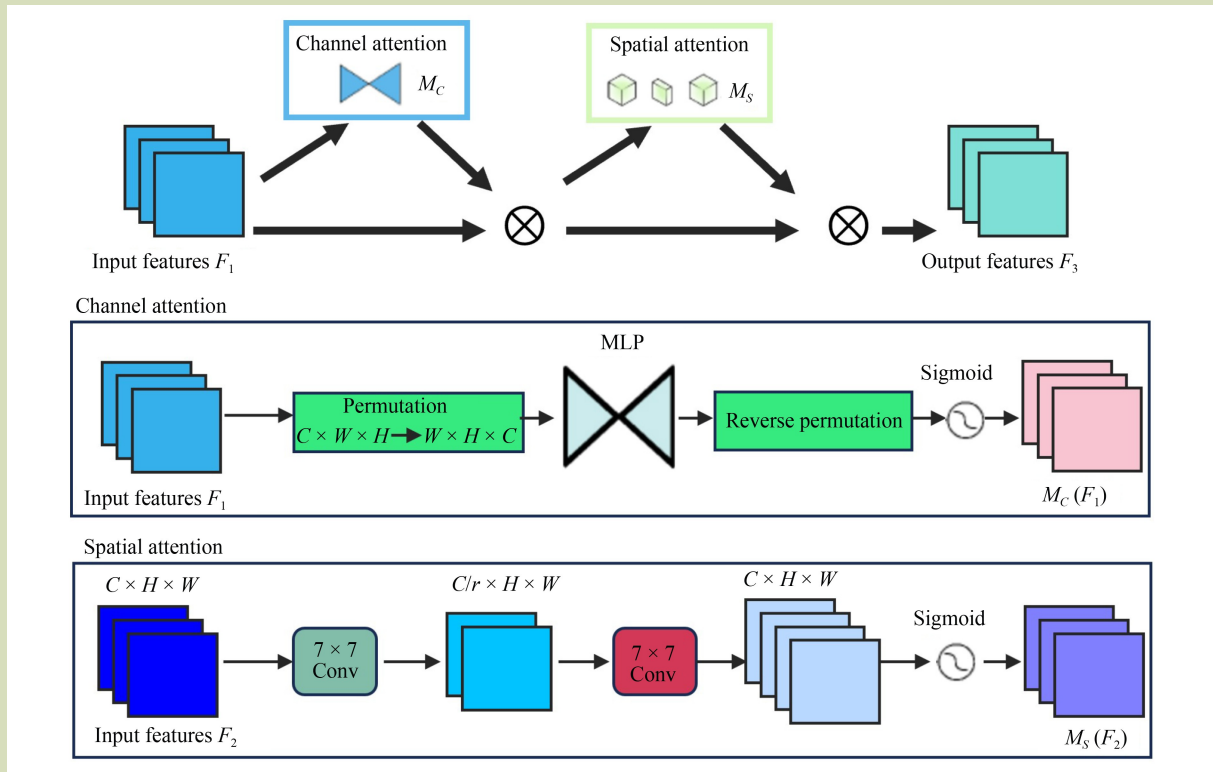


Fig. 4 Structure of the GAM module.

2.3.1.2 Convolution module improvement

To address the challenges associated with irregular canopy contours, elongated and curved inter-row gaps, and canopy occlusion, which can hinder conventional convolutional operations from accurately capturing geometric features along object boundaries, the original C3k2 module in the YOLOv11 backbone was replaced with the C3k2_DySnakeConv module, as illustrated in Fig. 5. The principal advantage of DySnakeConv lies in its serpentine sampling strategy, which follows the elongation direction of the target structure. Unlike the fixed-grid sampling employed by conventional convolution, DySnakeConv allows the convolutional kernel to adapt dynamically to the curved edges and elongated inter-row structures of *C. oleifera* canopies, thereby reducing feature misalignment caused by deviations between the receptive field and target geometry^[33]. In addition, its sub-pixel feature sampling strategy enlarges the effective receptive field without substantially increasing computational complexity, thereby improving feature representation in curved, discontinuous, and occluded regions.

2.3.1.3 YOLO model testing environment and evaluation metrics

The model training and testing were carried out on a computer system with the Windows 10 operating platform. In terms of hardware, the system specifications included an Intel® Xeon® Gold 6130 CPU (@ 2.10 GHz), Nvidia Quadro RTX 8000 GPU (48 GB memory), and 128 GB of RAM. Python 3.11 and the Pytorch 2.4.1 deep learning framework were used in the software environment, with CUDA 12.4 being used for GPU acceleration. The maximum number of training epochs was 300, while the batch size was set to 32. An early-stopping scheme was applied to reduce overfitting, and was implemented with 30 epochs of patience. The standard learning rate was 0.001.

The measurement of the model detection performance was conducted using the precision (P), recall (R), and mean average precision (mAP). The model deployability was measured based on the model size, while the real-time detection performance was assessed based on the average inference time per image (s).

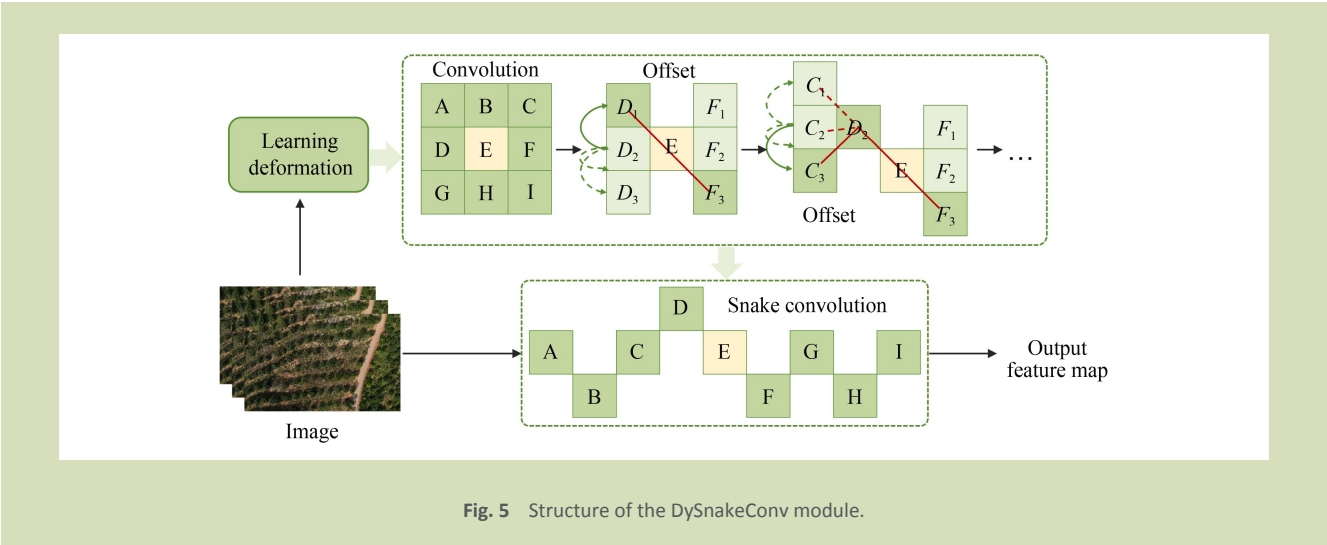


Fig. 5 Structure of the DySnakeConv module.

2.3.2 Navigation line detection

2.3.2.1 Navigation line generation method

To overcome the limitations of conventional approaches that directly fit the center points of detection bounding boxes, which may result in poor path continuity and inadequate robustness in the presence of curved crop rows or severe canopy occlusion, this study proposes a navigation path detection method that integrates clustering with spline-curve fitting. The detailed procedure is described below:

(1) Canopy center point extraction: For each predicted canopy bounding box $(x_{min}, y_{min}, x_{max}, y_{max})$, the center point coordinates (x_c, y_c) are calculated as:

$$x_c = \frac{x_{min} + x_{max}}{2}, y_c = \frac{y_{min} + y_{max}}{2} \quad (1)$$

(2) Row-line clustering and sorting based on K-Means: Considering the spatial distribution characteristics of *C. oleifera* rows, the K-Means algorithm is adopted to cluster the canopy center points. In actual orchard scenarios, the canopy centers exhibit more dispersed distribution and larger variance along the direction perpendicular to crop rows. Therefore, the coordinate direction with the maximum variance of center points is selected as the clustering feature. To realize automatic clustering for incomplete boundary rows and irregular planting blocks, a silhouette coefficient-based adaptive K-value optimization strategy is applied. After setting a reasonable search range for K, the optimal cluster number is automatically determined by maximizing the silhouette score. The resulting clusters were subsequently sorted according to their mean coordinates, ensuring that the order of the extracted

crop rows corresponded to their actual spatial arrangement in the orchard.

(3) Curve fitting and smoothing: To mitigate the effects of false detections and noise in natural orchard scenes, an outlier detection procedure was first introduced. The IQR method was employed to identify and remove points exhibiting excessive deviations from the main distribution. To avoid the limitations of fixed empirical thresholds in variable scenarios, the IQR multiplier factor and maximum removal ratio are treated as tunable hyperparameters and adaptively determined via grid search. Following outlier removal, cubic spline interpolation was applied to the remaining canopy center points to generate smooth row curves. The average curvature was subsequently calculated as an indicator of path smoothness, thereby enabling evaluation of the continuity and geometric smoothness of the fitted navigation paths.

(4) Navigation line generation: To satisfy the operational requirements of agricultural machinery performing inter-row operations, the centerline between two adjacent fitted crop-row curves was generated as the corresponding travel path. The principle of navigation-line generation is illustrated in Fig. 6.

2.3.2.2 Evaluation of navigation line extraction performance

To comprehensively evaluate the accuracy of the generated navigation paths, evaluation metrics were established at both global and local levels.

(1) At the global level, the Crop Row Detection Accuracy

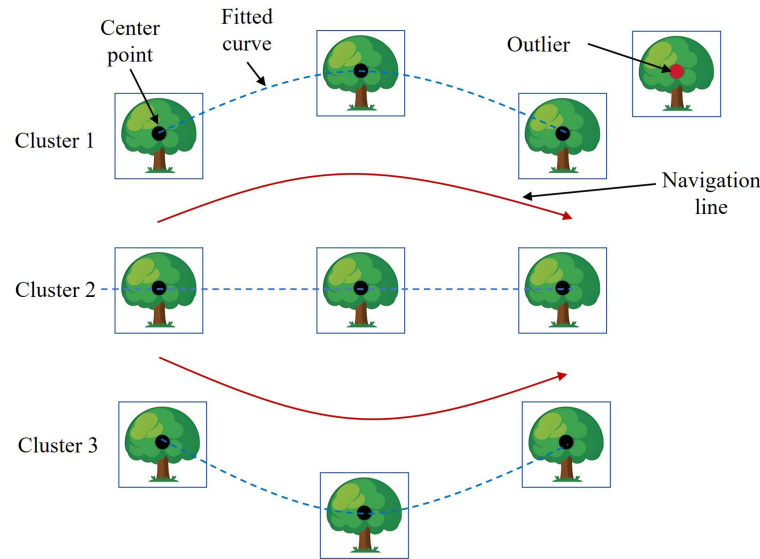


Fig. 6 Schematic diagram of navigation line generation.

(CRDA) was introduced as an evaluation metric^[34]. This metric calculates the matching score between the detected navigation lines and the manually annotated ground truth on a row-by-row basis, effectively reflecting the pixel-level consistency of the detection results.

Let H be the number of effective columns in the image, and let the detection algorithm output m navigation lines. For the v -th column of pixels ($v = 1, 2, \dots, H$), let the horizontal coordinate of the i -th detected navigation line be $u_{v,i}$, and the corresponding manually annotated ground truth horizontal coordinate be $u_{v,i}^*$.

First, the average row spacing d_v for the v -th row is calculated as:

$$d_v = \frac{1}{m-1} \sum_{i=1}^{m-1} (u_{v,i+1}^* - u_{v,i}^*) \quad (2)$$

This row spacing reflects the inherent distance between crop rows in the image and serves as the baseline for the subsequent deviation tolerance threshold.

To measure the matching degree between a detected point and the ground truth, the matching score function $s(u^*, u, d)$ is defined as follows:

$$s(u^*, u, d) = \max\left(1 - \left(\frac{u^* - u}{\sigma d}\right)^2, 0\right) \quad (3)$$

Where σ is a manually set tolerance coefficient. This function has the following properties: the score is 1 when the detected value exactly matches the ground truth; the score decays quadratically as the deviation increases when the deviation is less than σd ; and the score becomes zero when the deviation reaches or exceeds σd .

Based on the above matching score function, CRDA is defined as the average matching score over all effective rows and all navigation lines:

$$\text{CRDA} = \frac{1}{m \cdot H} \sum_{v=1}^H \sum_{i=1}^m s(u_{v,i}^*, u_{v,i}, d_v) \quad (4)$$

The CRDA ranges from 0 to 1. The closer the score is to 1, the higher the overall consistency between the detected navigation lines and the manually annotated ground truth.

(2) At the local level, the mean yaw angle and the mean lateral pixel distance were adopted as evaluation metrics^[22,35,36].

As shown in Fig. 7, let the slopes of the locally annotated navigation line and the predicted navigation line be k_1 and k_2 , respectively. The yaw angle θ was calculated as:

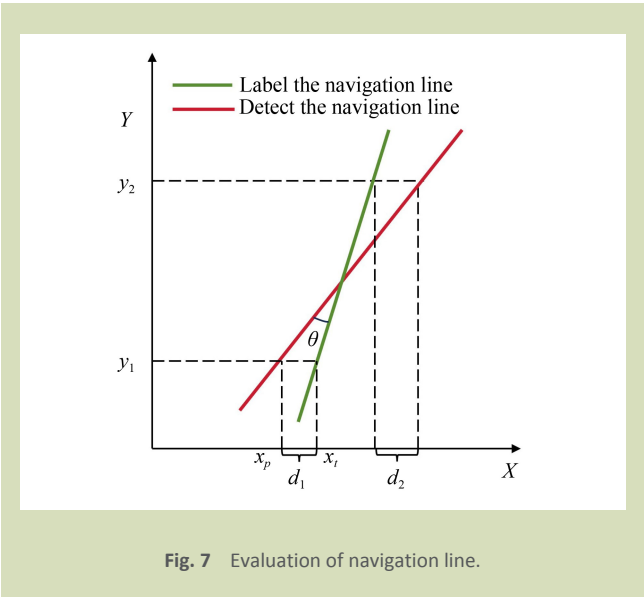


Fig. 7 Evaluation of navigation line.

$$\theta = \left| \arctan\left(\frac{1}{k_1}\right) - \arctan\left(\frac{1}{k_2}\right) \right| \tag{5}$$

After taking the absolute value, the mean yaw angle $\bar{\theta}$ was calculated across multiple images:

$$\bar{\theta} = \frac{1}{n} \sum_{i=1}^n \theta_i \tag{6}$$

The lateral pixel distance was defined as the absolute difference between the horizontal coordinates of the predicted navigation line and the annotated navigation line at the same vertical coordinate. Two reference points with a relatively large vertical separation were selected along the image, and the corresponding horizontal coordinate differences d_1 and d_2 were calculated. The mean lateral pixel distance $\bar{d}$ was then obtained as:

$$\bar{d} = \frac{|x_{p1} - x_{i1}| + |x_{p2} - x_{i2}|}{2} \tag{7}$$

Where x_p and x_i denote the horizontal coordinates of the predicted path and the annotated path, respectively, at the same vertical coordinate.

3 Results and discussion

3.1 Comparison with other models

To evaluate the effectiveness of the proposed DSConv-YOLOv11 model for *C. oleifera* tree canopy detection, comparative experiments were conducted against several representative lightweight YOLO models, including YOLOv5n, YOLOv8n, YOLOv10n, YOLOv11n, and YOLOv11s. Precision (P), recall (R), mAP, model size, and average inference time per image were selected as evaluation metrics. The comparative results are presented in Table 1.

As shown in Table 1, the proposed DSConv-YOLOv11 model demonstrated superior detection accuracy compared with the other models. It achieved a precision of 83.6%, which was 4.5%, 3.8%, 5.3%, 3.9%, and 2.4% higher than those of YOLOv5n, YOLOv8n, YOLOv10n, YOLOv11n, and YOLOv11s, respectively. Its recall reached 79.3%, representing improvements of 4.0%, 2.5%, 3.4%, 4.1%, and 1.0%, respectively. The mAP reached 85.3%, exceeding the corresponding values of the five baseline models by 4.7%, 3.5%, 4.9%, 5.2%, and 2.6%, respectively. This demonstrates that the proposed DSConv-YOLOv11 can better recognize and locate *C. oleifera* canopies and reduce both missed detections and false positives. In terms of overall model size and real-time performance, the weight size of DSConv-YOLOv11 is 5.07 MB. Equipped with the GAM attention mechanism and DySnakeConv module, the model improves overall localization and recognition accuracy while remaining lightweight to meet

Table 1 Comparative experimental results with YOLO models					
Model	P (%)	R (%)	mAP (%)	Weight (MB)	Inference time (ms)
YOLOv5n	79.1	75.3	80.6	2.53	306.87
YOLOv8n	79.8	76.8	81.8	2.60	300.25
YOLOv10n	78.3	75.9	80.4	2.69	270.01
YOLOv11n	79.7	75.2	80.1	2.58	296.85
YOLOv11s	81.2	78.3	82.7	9.43	338.63
DSConv-YOLOv11	83.6	79.3	85.3	5.07	303.33

the deployment requirements of low-altitude remote sensing sensors. The average inference time is 303.33 ms per image, which is comparable to YOLOv5n and YOLOv11n and does not bring significantly increased inference latency. Overall, the proposed DSConv-YOLOv11 achieves balanced detection performance and efficiency.

Figure 8 shows the changes in precision and recall of models during training. As observed, DSConv-YOLOv11 achieves higher precision and recall throughout training compared with the baseline models YOLOv5n, YOLOv8n, YOLOv10n, YOLOv11n, and YOLOv11s. In addition, the proposed DSConv-YOLOv11 achieves a faster convergence speed, higher

final accuracy and lower training curve fluctuations, indicating stable training without obvious oscillation. Overall, these results confirm that the proposed structural modifications enhance feature extraction and learning capabilities to realize more accurate and stable canopy detection.

To further evaluate the overall performance of DSConv-YOLOv11, three representative crop detection algorithms reported in existing studies, namely ASPPF-YOLOv8s^[16], Maize-YOLO^[18] and ResC-UNet^[20], were selected for comparative experiments. As summarized in Table 2, DSConv-YOLOv11 achieves higher precision and recall than ASPPF-YOLOv8s and Maize-YOLO. Notably, these two models are

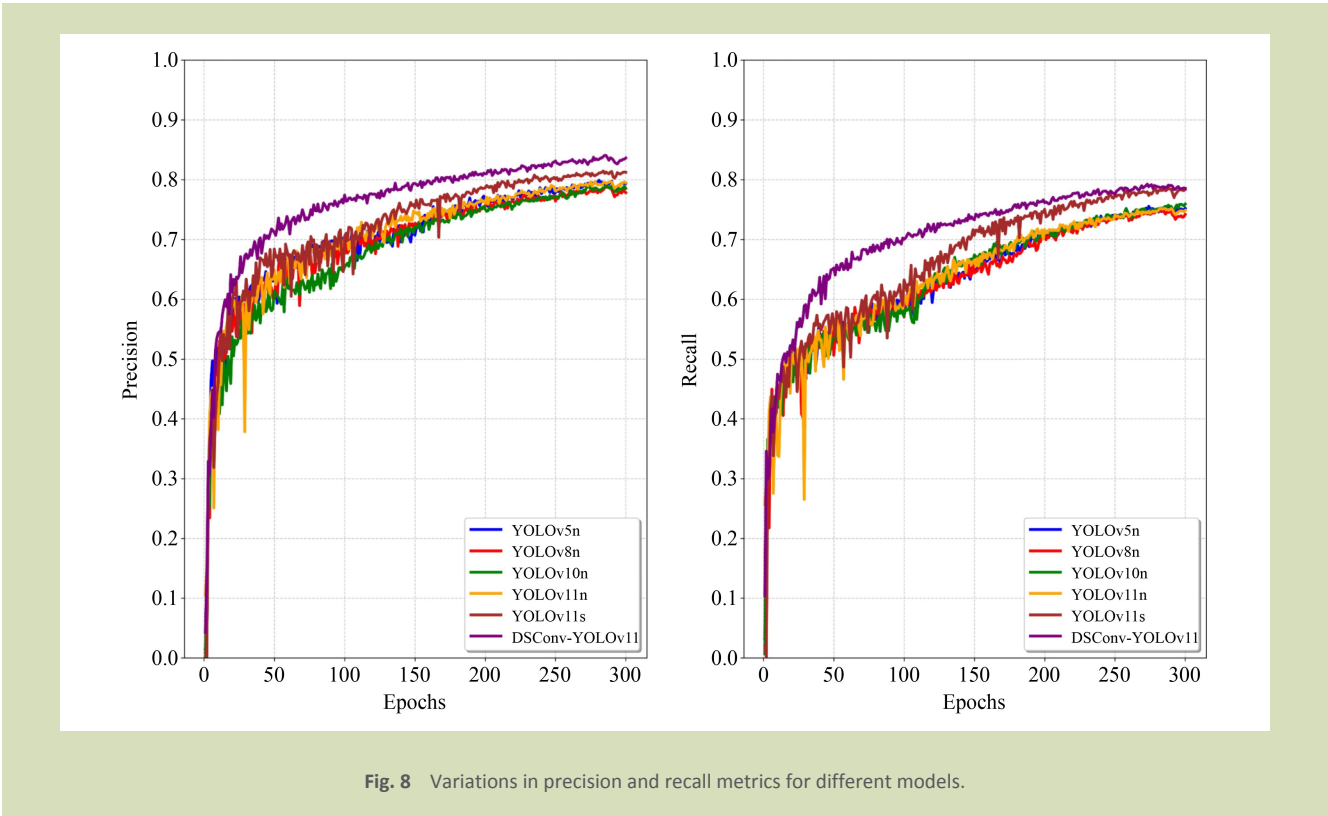


Fig. 8 Variations in precision and recall metrics for different models.

Table 2 Comparative experimental results with other models					
Model	P (%)	R (%)	mAP (%)	Weight (MB)	Inference time (ms)
ASPPF-YOLOv8s	82.4	78.1	84.2	20.49	271.00
Maize-YOLO	83.1	78.6	84.8	12.14	86.06
ResC-UNet	83.5	79.4	86.0	25.96	434.70
DSConv-YOLOv11	83.6	79.3	85.3	5.07	303.33

specially developed for crop detection on flat farmlands. Although ResC-UNet achieves marginally higher mAP via a pixel-level semantic segmentation strategy, its model size reaches 25.96 MB, leading to markedly inferior inference efficiency. By contrast, DSConv-YOLOv11 requires only 5.07 MB of storage space and delivers competitive detection performance together with favorable inference speed.

In summary, the proposed DSConv-YOLOv11 model integrates GAM to enhance feature focusing and DySnakeConv to improve geometric feature perception, thereby achieving a favorable balance between detection accuracy and real-time performance for inter-row navigation line detection in *C. oleifera* orchards.

3.2 Comparative experiments with different module replacements

3.2.1 Comparative experiments with different feature fusion mechanisms

To verify the effectiveness of GAM for feature fusion in *C. oleifera* tree canopy detection, comparative experiments were performed by substituting GAM with CBAM^[37], EMA^[38], and MSCA^[39]. All other experimental conditions,

including the backbone network and convolutional architecture, remained unchanged. As shown in Table 3, replacing the original module with CBAM, EMA and GAM improved model performance, among which GAM brought the most prominent performance gain. In contrast, replacing GAM with MSCA reduced detection accuracy, implying that MSCA exhibits poor adaptability to the complex features within *C. oleifera* orchard scenarios. With the integration of GAM, the model achieved a precision of 82.5%, a recall of 78.6%, and a mAP of 83.6%, representing improvements of 2.8%, 3.4% and 3.5%, respectively, compared with the baseline model. These metrics also outperformed those obtained using CBAM, EMA and MSCA. Moreover, embedding GAM introduced negligible computational overhead, with the model size and inference time remaining within acceptable ranges.

3.2.2 Comparative experiments with different convolution module replacements

To validate the rationality of replacing the original C3k2 module with the C3k2_DySnakeConv module, comparative experiments were conducted using three alternative convolutional modules. The quantitative results are presented in Table 4. Compared with the baseline model and the models embedded with DynamicConv^[40], PConv^[41], and RFB^[42], the proposed model with C3k2_DySnakeConv achieved mAP

Table 3 Comparative study of different feature fusion mechanisms					
Model	P (%)	R (%)	mAP (%)	Weight (MB)	Inference time (ms)
YOLOv11n	79.7	75.2	80.1	2.58	296.85
YOLOv11-CBAM	79.8	75.3	80.6	2.65	284.32
YOLOv11-EMA	79.9	76.2	80.8	2.63	295.43
YOLOv11-MSCA	79.2	73.7	79.5	2.74	304.31
YOLOv11-GAM	82.5	78.6	83.6	5.06	296.81

Table 4 Comparative study of different convolutional variants					
Model	P (%)	R (%)	mAP (%)	Weight (MB)	Inference time (ms)
YOLOv11n	79.7	75.2	80.1	2.58	296.85
YOLOv11-DynamicConv	78.8	75.6	80.0	4.25	293.65
YOLOv11-PConv	79.8	75.7	80.4	2.61	305.20
YOLOv11-RFB	78.6	73.8	78.9	2.82	301.23
YOLOv11-C3k2_DySnakeConv	80.2	76.3	81.7	3.43	318.40

improvements of 1.6%, 1.7%, 1.3%, and 2.8%, respectively. Meanwhile, both precision and recall were superior to those of the baseline model. These findings demonstrate that the replacement of the standard C3k2 module with C3k2_DySnakeConv effectively enhances the model’s ability to capture fine-grained edge features and inter-row structural characteristics of *C. oleifera* tree canopies, thereby improving detection accuracy and stability in complex orchard scenarios.

3.3 Model ablation experiments and analysis

To analyze the effects of the GAM attention mechanism and the C3k2_DySnakeConv dynamic serpentine convolution on the detection metrics of *C. oleifera* canopy detection models, we conducted ablation experiments. Each group was independently tested five times with all other hyperparameters kept identical, and the results are reported as the mean ± standard deviation of mAP and inference time. The experimental results show that the sole integration of GAM increased the mAP from 80.1% to 83.6%, corresponding to a 3.5% improvement. The individual adoption of C3k2_DySnakeConv improved the mAP by 1.6%. The simultaneous integration of both modules further increased the mAP to 85.3%, with a maximum performance improvement of 5.2%. As presented in Table 5, the mAP standard deviation across all experimental groups was lower than 0.3%, which is considerably smaller than the performance gains achieved by the proposed modules. This indicates that the performance improvements stem mainly from the structural modifications to the baseline model, rather than random fluctuations during training. Furthermore, merely introducing C3k2_DySnakeConv led to increased inference latency, whereas the collaborative application of GAM and C3k2_DySnakeConv achieves a favorable balance between detection accuracy and real-time deployment requirements.

3.4 Navigation line extraction test

To evaluate the effectiveness of the proposed navigation line generation method, experiments on navigation line extraction and accuracy assessment were conducted using aerial images of a *C. oleifera* orchard.

Initially, we adopted DSConv-YOLOv11 to detect canopies of *C. oleifera* trees and distinguish inter-row areas from tree regions. The detection results are presented in Fig. 9(a). After detection, we extracted the center points of the detected tree canopies. Based on their spatial distribution, we further performed clustering, outlier elimination and spline fitting to generate curves characterizing the distribution of *C. oleifera* rows, as illustrated in Fig. 9(b). Inter-row navigation lines were subsequently constructed using the fitted data (Fig. 9(c)). Figure 9(d) shows the reference navigation line generated from manually labelled data. By comparing Fig. 9(c,d), the navigation lines automatically produced by our method are well aligned with the reference trajectory. The generated navigation lines exhibit favorable continuity and smoothness, particularly in curved areas suffering from canopy occlusion and overlapping tree rows. Only slight deviations can be observed for sharply curved segments.

To quantitatively evaluate the quality of navigation line detection, we adopted the CRDA metric proposed in Yin et al.[17]. CRDA ranges from 0 to 1, where a value closer to 1 indicates higher detection accuracy. To investigate how the evaluation results, depend on the tolerance, we conducted a sensitivity analysis with tolerance coefficients varying from 0.05 to 0.25. As shown in Fig. 10, the mean CRDA values of all three orchards increase monotonically as the tolerance coefficient rises. A larger tolerance coefficient loosens the matching criterion for performance evaluation, which naturally leads to higher CRDA values. It is worth noting that CRDA is

Table 5 Ablation experiment results

Test No.	GAM	C3k2_DySnakeConv	mAP (%)	Weight (MB)	Inference time (ms)
1	–	–	80.1 ± 0.21	2.58	296.85 ± 1.12
2	√	–	83.6 ± 0.25	5.06	296.81 ± 1.08
3	–	√	81.7 ± 0.23	3.43	318.40 ± 1.35
4	√	√	85.3 ± 0.27	5.07	303.33 ± 1.20

Note: “–” indicates that the corresponding module is not adopted in the model; “√” indicates that the corresponding module is adopted in the model.

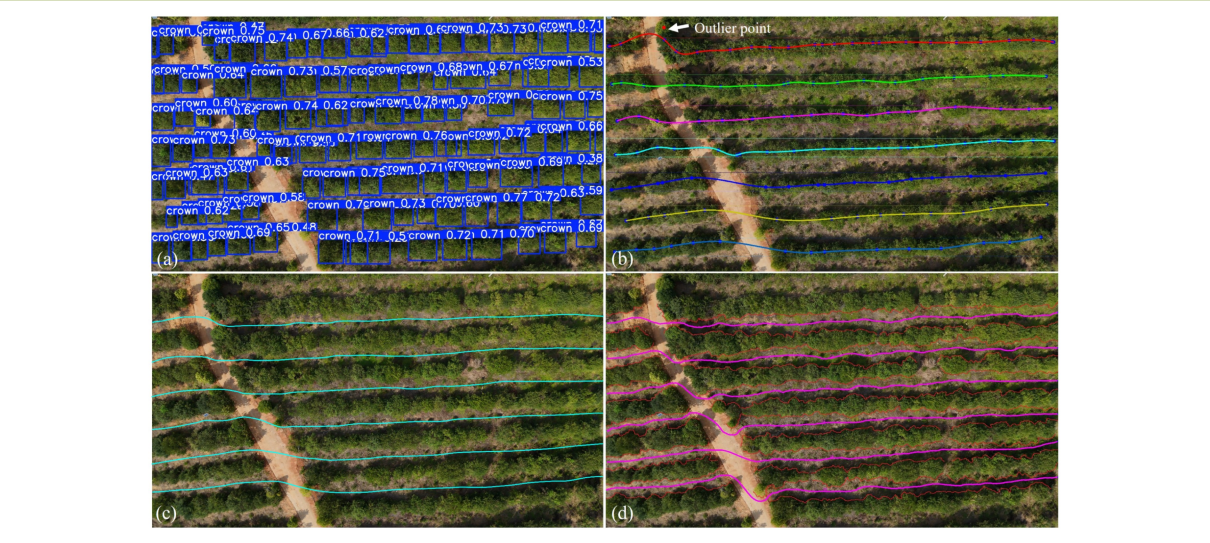


Fig. 9 Navigation line generation. (a) *C. oleifera* canopy detection results; (b) *C. oleifera* distribution curve; (c) navigation lines; (d) annotated navigation lines.

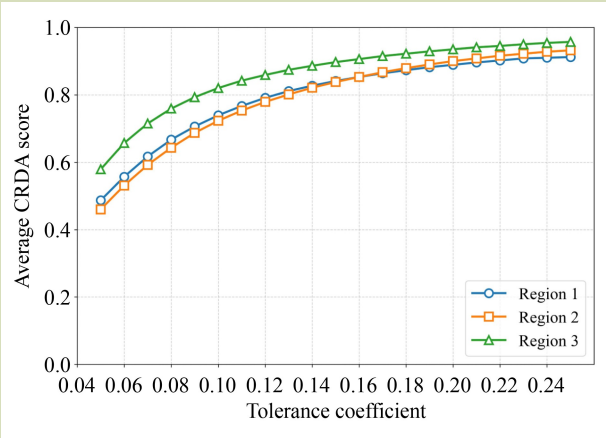


Fig. 10 Average CRDA scores under different tolerance coefficients.

not an absolute quantitative metric; its value depends on the observer bound parameter. In this work, the tolerance coefficient σ was selected following the suggestion in the literature^[34], where the allowable deviation threshold is set to 10% of the crop row spacing. To intuitively visualize the detection quality along image columns, we plotted a heatmap, in which color intensity reflects the detection quality at each corresponding crop row.

As shown in Fig. 11, the heatmap exhibits an overall color distribution characterized by darker tones in the center and lighter tones towards both ends, indicating that the navigation-line scores in the central region of the image are generally higher than those in the edge regions. This finding suggests that detection conditions are more favorable in the central area, whereas the edge regions are adversely affected by incomplete canopy occlusion and lens distortion, which have a non-negligible impact on detection accuracy. Furthermore, the distribution of low-score regions is broadly consistent with the locations of internal roadways within the field blocks. This occurs because roadways disrupt the continuity of canopy center points across adjacent crop rows, resulting in segments with insufficient valid reference points. Consequently, the linking algorithm tends to generate smoother approximations of the navigation paths, causing the resulting trajectories to deviate from the ideal routes and thereby producing relatively lower scores in these regions.

As shown in the bar chart, the CRDA scores of individual navigation lines range from 0.4479 to 0.9042. Among the three field blocks, Block 3 achieved the best overall performance, with a mean CRDA score of 0.8317. Navigation Line 2 in this block achieved the highest individual score of 0.9042, indicating a high degree of agreement between the detected navigation line and the manually annotated ground truth and

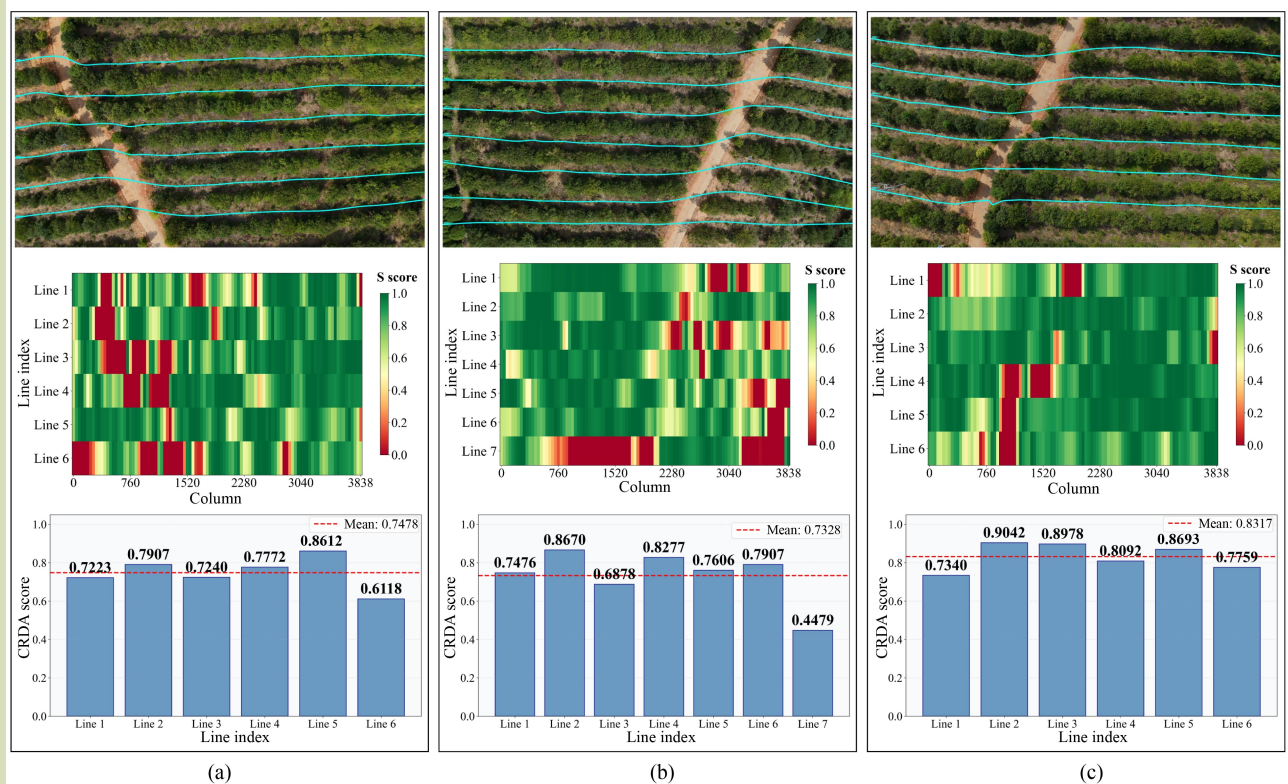


Fig. 11 CRDA scores of different navigation lines. (a) Region 1; (b) region 2; (c) region 3.

demonstrating the effectiveness of the proposed method under favorable conditions. Conversely, Block 2 showed a very low mean of 0.7328 with large dispersion within individual navigation lines, which might relate to the uneven distribution of *C. oleifera* trees on the bottom right side of the field. Notably, the lowest CRDA score was recorded for Navigation Line 7 in Block 2, primarily because only part of the canopy was visible in this row, resulting in deviations during canopy center-point detection and subsequent computation.

Overall, the mean CRDA scores for all three field blocks exceeded 0.73, with the highest-performing block approaching 0.83, while the majority of individual navigation lines achieved scores above 0.75. These results demonstrate the reliability of the proposed method across most field conditions. However, image-edge regions and areas with discontinuous crop rows remain the main factors limiting overall detection accuracy. These findings provide useful guidance for further optimization of the proposed navigation-line detection algorithm.

For the local accuracy evaluation, heading deviation and lateral deviation were selected as the primary evaluation metrics. To assess the local fitting quality of the generated navigation lines across different path segments, a sliding-window analysis was employed. The length of the window was set to 40 pixels, with a step size of one pixel. The navigation path was then sampled in a continuous manner. For every sampled window within every field block, the average of the deviations between the navigation lines created and the manual reference annotation paths were determined. An example of the deviations is shown in Fig. 12. As can be seen in the figures, deviations were markedly higher between the sampling points which corresponded to sections of roadways, as the roadway presence interrupted the visual continuity of the crop rows. In spite of these relatively high values of deviations from the aforementioned two-dimensional space, the maximum heading deviation was below 17.5° and the maximum lateral deviation was below 31 pixels. Out of the total number of windows sampled, less than 5% had segments experiencing such extreme deviations. Considering the low operating speed of agricultural

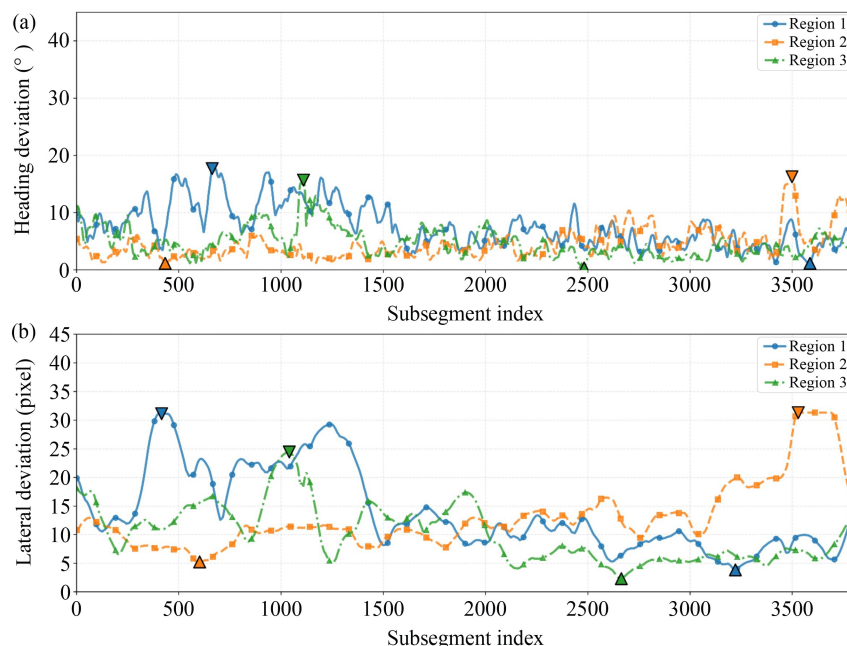


Fig. 12 Line chart of heading (a) and lateral (b) deviations.

navigation machinery, such deviations are acceptable for navigation tasks. In the remaining segments, both deviation measures were generally lower than the overall average, demonstrating good consistency between the generated and reference paths.

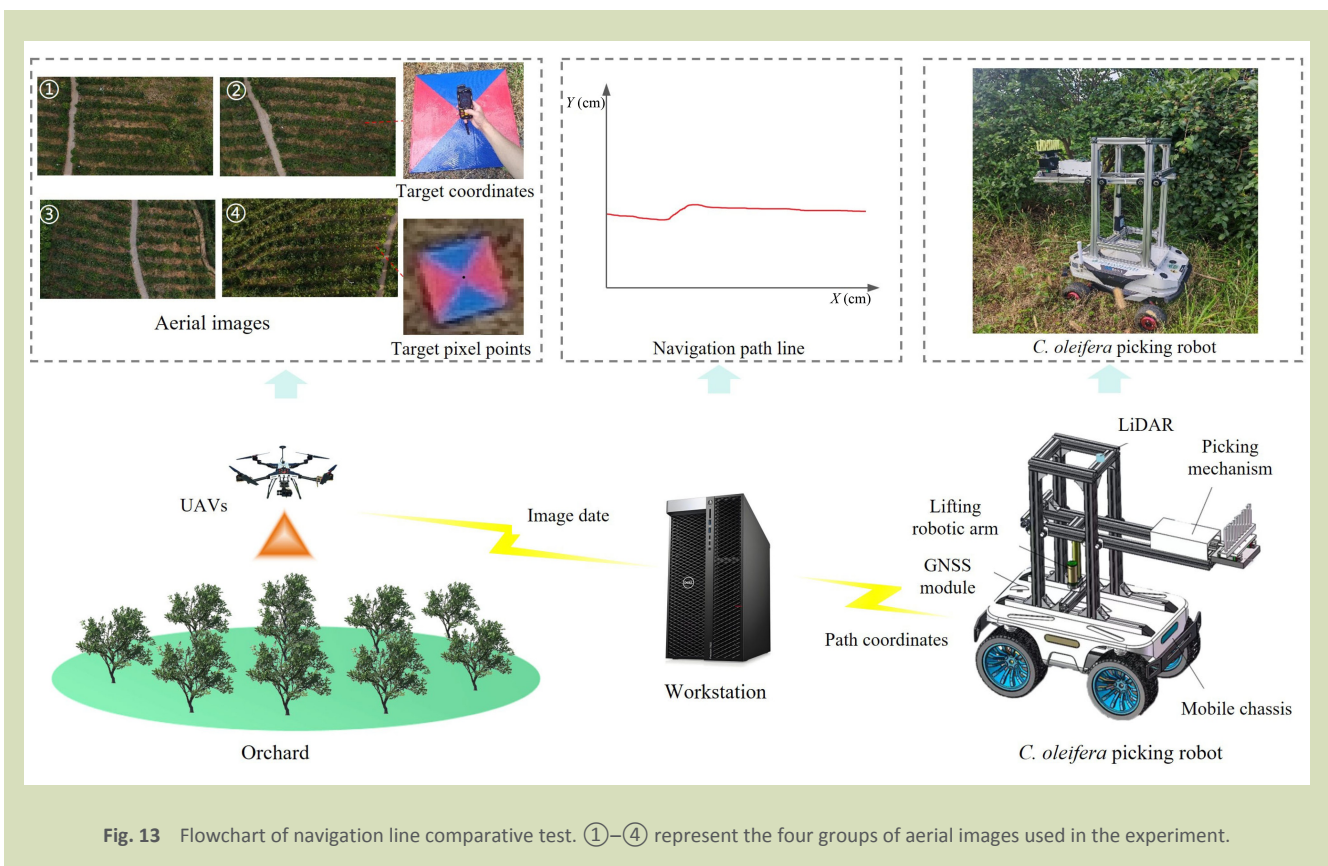
Across the three field blocks, the heading deviation ranged from 0.3315° to 17.2265° , with a mean of 5.25° , whereas the lateral deviation ranged from 2.28 to 30.76 pixels, with a mean of 11.62 pixels. These results indicate that although local fluctuations occurred in some segments because of crop-row discontinuities caused by roadways, the overall deviations remained within an acceptable range^[14]. Overall, the local deviation fluctuations of the navigation lines generated by the proposed method were limited and controllable, satisfying the accuracy requirements for inter-row navigation in *C. oleifera* orchards.

3.5 Comparative test of navigation lines

To evaluate the navigation accuracy of the proposed algorithm under practical field operating conditions, comparative experiments were conducted in this section. The comparison

focused on the global navigation trajectories generated from low-altitude aerial images and those produced by the onboard perception-based navigation system of the *C. oleifera* picking robot. The experiments were conducted at a *C. oleifera* planting base in Yongzhou City, Hunan Province, where the environmental conditions were broadly consistent with those encountered during dataset acquisition.

The experimental procedure is illustrated in Fig. 13. First, eight positioning targets were uniformly distributed throughout the test area, and the high-precision geodetic coordinates of each target were obtained using a GNSS positioning device. Meanwhile, complete aerial images of the experimental plot were acquired using a UAV and transmitted to a ground workstation. The inter-row reference navigation lines were then extracted using the proposed DSConv-YOLOv11-based navigation-line detection algorithm. By matching the two-dimensional pixel coordinates of the positioning targets with their corresponding three-dimensional geographic coordinates, a homography transformation matrix between the image plane and the ground plane was established. This transformation enabled the pixel-level navigation paths to be mapped to planar geographic coordinates with centimetre-level positional accuracy.



During the field test, the *C. oleifera* picking robot was operated in fully autonomous navigation mode and travelled along the inter-row regions. Path coordinates were sampled at uniform 1-m intervals, yielding a total of 100 sampling points. These measurements were subsequently used for quantitative deviation analysis and comparison between the two navigation-path schemes.

The absolute deviations between the corresponding sampling points of the two path sets were calculated, and the resulting deviation distributions for the four test groups are presented as boxplots in Fig. 14. The median path deviations for the four groups were 37.7, 30.8, 27.6, and 38.8 cm, respectively, with corresponding mean deviations of 51.8, 43.2, 45.2, and 55.7 cm. The experimental results indicate that the navigation paths generated from UAV aerial imagery exhibited deviation magnitudes comparable to those of the onboard navigation paths of the *C. oleifera* picking robot. Thus, the trajectories generated by the proposed algorithm can generally satisfy the operational requirements of autonomous inter-row navigation in *C. oleifera* orchards.

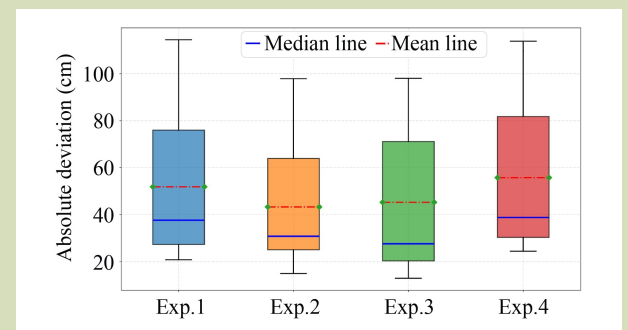


Fig. 14 Comparison results with the trajectory of the *C. oleifera* picking robot.

Nevertheless, slight fluctuations in deviation were observed across all four groups, with the maximum absolute deviation reaching approximately 110 cm. These deviations may primarily be attributed to trajectory offsets of the picking robot and the neglect of terrain undulations during the planar perspective transformation. In future work, terrain elevation data could be incorporated into the coordinate transformation

model to improve mapping accuracy. In addition, real-time onboard visual information could be integrated to dynamically adjust the navigation paths, thereby further improving trajectory consistency and positioning accuracy.

4 Discussion

The purpose of this paper is to detect inter-row navigation lines in *C. oleifera* orchards. Major challenges exist in this scenario, including uneven canopy distribution, heavy shadows, discontinuous crop rows, and the limited field of view of ground vision systems, all of which degrade detection reliability. To address these problems, this study proposes a UAV-based inter-row navigation line detection method. Specifically, the method consists of the newly developed DSConv-YOLOv11 detector together with a navigation line generation strategy combining K-Means clustering, IQR-based outlier elimination and cubic spline fitting.

Experimental results show that canopy detection accuracy is positively correlated with navigation line quality. In particular, the proposed DSConv-YOLOv11 detector achieves superior mAP compared with YOLOv11n. The mAP of YOLOv11n reaches 80.1%, while the proposed model improves this metric to 85.3%. Correspondingly, precision and recall increase by 3.9% and 4.1%, respectively. Owing to more accurate canopy center coordinates, the improved detection performance further enhances navigation line extraction accuracy. Although GAM and DySnakeConv enlarge the model size to approximately 5.07 MB and yield an average inference time of around 303 ms per aerial image, the method achieves a favorable trade-off between detection accuracy and inference efficiency.

Ablation experiments are conducted to verify the independent contribution of each module. Incorporating the GAM module improves mAP by 3.5%. This indicates that enhanced global channel-spatial interactions can suppress background interference and strengthen feature representation for *C. oleifera* canopies. Using only the C3k2_DySnakeConv module raises mAP by 1.6%, which verifies that serpentine convolution effectively extracts geometric features under uneven illumination and irregular canopy edge conditions.

Navigation experiments confirm that merely boosting object detection accuracy is insufficient. The proposed post-

processing pipeline enhances the continuity and smoothness of navigation lines under real orchard conditions. The CRDA value exceeds 0.73 across all experimental plots. The average heading deviation and lateral deviation of generated routes are 5.25° and 11.62 pixels, respectively. Larger deviations mainly occur near internal roads and canopy breakpoints, where missing trees break the spatial continuity of planting rows. Nevertheless, the generated navigation paths preserve global continuity, which validates the effectiveness of the proposed fitting strategy. Visual inspection further verifies that the proposed method achieves strong robustness in orchard environments.

Recent studies have developed a range of vision-based navigation approaches for agricultural robots. Diao et al.^[16] extracted maize navigation lines using an improved YOLOv8s detector and achieved high detection accuracy in relatively regular farmland environments. Ying et al.^[22] integrated YOLOv8-Seg with edge extraction to generate navigation paths for strawberry-harvesting robots, whereas Shi et al.^[26] employed UAV imagery to perform global path planning in orchards based on fruit-tree localisation. Ye et al.^[43] further demonstrated the importance of path-smoothness optimisation for mobile robot navigation in complex orchard environments through the development of the KBQ-RRT* sampling-based path planner. Unlike these existing methods, the proposed framework is specially tailored for *C. oleifera* orchards characterized by curved planting layouts, irregular canopy shapes and complex mountainous terrain, which brings greater difficulties to navigation tasks. The proposed method relies neither solely on object detection nor semantic segmentation. Instead, it combines lightweight object detection with geometric clustering and spline fitting to generate continuous navigation paths even when canopies are discontinuously distributed. The mAP of 85.3% and mean CRDA above 0.73 reveal that the proposed method achieves competitive detection accuracy and exhibits favorable navigation performance under highly complex orchard conditions.

From an engineering perspective, the proposed framework possesses prominent practical application value. UAVs can rapidly acquire global environmental information of orchards, and the lightweight DSConv-YOLOv11 supports efficient inference on workstations. The navigation lines extracted by this framework can serve directly as reference trajectories for autonomous tractors, spraying robots, harvesting robots and other unmanned equipment operating in *C. oleifera* orchards.

Despite the promising experimental results, several limitations remain to be addressed. First, all experiments were carried out under favorable weather conditions. The model robustness under heavy rainfall, fog and drastic seasonal illumination variations requires further investigation. Second, the current K-Means clustering algorithm exhibits limited adaptability to orchard scenes featuring irregular trees and discontinuous crop rows. Third, navigation accuracy was evaluated against manually labeled reference paths and robot onboard navigation trajectories, while closed-loop navigation experiments with real-time robot control have not yet been completed. Future research will integrate the proposed navigation framework with ground autonomous vehicles and path-tracking controllers. Meanwhile, the dataset will be expanded to cover diverse growth stages, orchard types and environmental conditions. Self-supervised learning strategies will also be explored to improve the robustness and generalization of the UAV-based orchard navigation system.

5 Conclusions

This study proposes a UAV-based navigation line detection method built upon DSConv-YOLOv11 to address the issues of limited global environmental perception and insufficient detection accuracy for *C. oleifera* canopies during inter-row navigation in *C. oleifera* orchards. The main conclusions are summarized as follows:

(1) The DSConv-YOLOv11 model embedded with the GAM attention mechanism and C3k2_DySnakeConv convolution module achieves notable performance improvements in *C. oleifera* canopy detection. Ablation experiments demonstrate that these improved modules jointly increase the mAP to 85.3%, with precision and recall reaching 83.6% and 79.3%, respectively. Compared with YOLOv11n, the three metrics are improved by 3.9%, 4.1% and 5.2%. Furthermore, the model size is only 5.07 MB, and the average inference time is 303.33 ms per image. These characteristics enable a favorable trade-off between detection accuracy and computational efficiency.

(2) The navigation line generation strategy proposed in this work integrates K-Means clustering, outlier removal and cubic spline fitting, which enables the generation of continuous and smooth inter-row navigation lines even under crop row shadow interference and segmentation discontinuities. Taking manually labelled trajectories as references, the navigation lines generated by our method yield an average CRDA above 0.73; the mean heading deviation and mean lateral deviation are 5.25° and 11.62 pixels, respectively. Field tests conducted using a UAV-assisted *C. oleifera* harvesting robot further verify that the overall deviation between the UAV-generated navigation lines and actual travelling trajectories remains relatively small. This result confirms that the extracted navigation lines can serve as reliable reference paths for autonomous agricultural ground vehicles.

Acknowledgements

This work was financially supported by the Natural Science Foundation of Hunan Province, China (2024JJ622) and the Scientific Research Project of the Hunan Provincial Department of Education, China (25A0593, 25C0592, 23B0761).

Compliance with ethics guidelines

Linhui Wang, Xiaowen Liu, Lizhi Chen, Yonghong Tan, Bin Qing, Yu Yang, and Mengyu Zeng declare that they have no conflicts of interest or financial conflicts to disclose. This article does not contain any studies with human or animal subjects performed by any of the authors.

REFERENCES

1. Campoverde-Molina M, Luján-Mora S. Cybersecurity in smart agriculture: a systematic literature review. *Computers & Security*, 2025, **150**: 104284
2. Qu J W, Zhang Z, Qin Z Y, Guo K Q, Li D. Applications of autonomous navigation technologies for unmanned agricultural tractors: a review. *Machines*, 2024, **12**(4): 218
3. Wang Q Y, He J, Lu C Y, Wang C, Lin H, Yang H Y, Li H, Wu Z Y. Modelling and control methods in path tracking control

- for autonomous agricultural vehicles: a review of state of the art and challenges. *Applied Sciences*, 2023, **13**(12): 7155
4. Yin X, Du J, Noguchi N, Yang T X, Jin C Q. Development of autonomous navigation system for rice transplanter. *International Journal of Agricultural and Biological Engineering*, 2018, **11**(6): 89–94
5. Kanagasingham S, Ekpanyapong M, Chaihan R. Integrating machine vision-based row guidance with GPS and compass-based routing to achieve autonomous navigation for a rice field weeding robot. *Precision Agriculture*, 2020, **21**(4): 831–855
6. Chen Y T, Fan J L, Li C L, Li B L, Wei Z B, Wang J. Research on orchard navigation path planning based on 3D LiDAR SLAM considering terrain roughness. *Journal of Field Robotics*, 2026, **43**(7): 4696–4719
7. Pan Y Q, Hu K W, Cao H, Kang H W, Wang X. A novel perception and semantic mapping method for robot autonomy in orchards. *Computers and Electronics in Agriculture*, 2024, **219**: 108769
8. Firkat E, An F, Peng B, Zhang J L, Mijit T, Ahat A, Zhu J H, Hamdulla A. FGSeg: field-ground segmentation for agricultural robot based on LiDAR. *Computers and Electronics in Agriculture*, 2023, **211**: 107965
9. de Silva R, Cielniak G, Wang G, Gao J F. Deep learning-based crop row detection for infield navigation of agri-robots. *Journal of Field Robotics*, 2024, **41**(7): 2299–2321
10. Li H W, Lai X D, Mo Y M, He D Q, Wu T. Pixel-wise navigation line extraction of cross-growth-stage seedlings in complex sugarcane fields and extension to corn and rice. *Frontiers in Plant Science*, 2025, **15**: 1499896
11. Liu Z P, Wu X L, Gao W W, Ning F P, Zhang W Y, He J, Liu M H, Chen X F, Fang P. Research on a vision navigation system for rice transplanters based on machine vision fused with GNSS. *Biosystems Engineering*, 2026, **262**: 104354
12. Zhang S, Liu Y, Xiong K, Tian Y H, Du Y F, Zhu Z X, Du M W, Zhai Z Q. A review of vision-based crop row detection method: focusing on field ground autonomous navigation operations. *Computers and Electronics in Agriculture*, 2024, **222**: 109086
13. Bai Y H, Zhang B H, Xu N M, Zhou J, Shi J Y, Diao Z H. Vision-based navigation and guidance for agricultural autonomous vehicles and robots: a review. *Computers and Electronics in Agriculture*, 2023, **205**: 107584
14. Wei P, Ragbir P, Vougioukas S G, Kong Z D. Vision-based navigation of unmanned aerial vehicles in orchards: an imitation learning approach. *Computers and Electronics in Agriculture*, 2025, **238**: 110802
15. Bai Y F, Chen Z R, Liang P P, Song B, Cheng E K. CurveFormer++: 3D lane detection by curve propagation with temporal curve queries and attention. *IEEE Transactions on Intelligent Transportation Systems*, 2025, **26**(6): 7909–7920
16. Diao Z H, Guo P L, Zhang B H, Zhang D Y, Yan J N, He Z D, Zhao S N, Zhao C J, Zhang J C. Navigation line extraction algorithm for corn spraying robot based on improved YOLOv8s network. *Computers and Electronics in Agriculture*, 2023, **212**: 108049
17. Yin Y X, Zeng W, Wang H, Zhang C T, Meng Z J. A corn seedling row detection method for fields based on YOLO-AP. *Computers and Electronics in Agriculture*, 2026, **248**: 111747
18. Yang Y, Zhou Y, Yue X, Zhang G, Wen X, Ma B, Xu L Y, Chen L Q. Real-time detection of crop rows in maize fields based on autonomous extraction of ROI. *Expert Systems with Applications*, 2023, **213**: 118826
19. Zheng S X, Yang X P, Gao P, Guo Q W, Zhang J H, Chen S H, Tang Y C. YOLOv11-IMP: anchor-free multiscale detection model for accurate grape yield estimation in precision viticulture. *Agronomy*, 2026, **16**(3): 370
20. Zhang X Y, Wang Q J, Wang X H, Li H W, He J, Lu C Y, Yang Y, Jiang S. Automated detection of crop-row lines and measurement of maize width for boom spraying. *Computers and Electronics in Agriculture*, 2023, **215**: 108406
21. Li D F, Li B L, Long S F, Feng H Q, Xi T, Kang S, Wang J. Rice seedling row detection based on morphological anchor points of rice stems. *Biosystems Engineering*, 2023, **226**: 71–85
22. Ying Q K, Cheng H C, Ma Z H, Du X Q. Ridge visual navigation control method for ground-planted strawberry picking robots based on YOLO v8-seg algorithm. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, **55**(S1): 9–17 (in Chinese)
23. Ruan Z W, Chang P H, Cui S Q, Luo J Q, Gao R, Su Z B. A precise crop row detection algorithm in complex farmland for unmanned agricultural machines. *Biosystems Engineering*, 2023, **232**: 1–12
24. Cao M Y, Tang F F, Ji P, Ma F Y. Improved real-time semantic segmentation network model for crop vision navigation line detection. *Frontiers in Plant Science*, 2022, **13**: 898131
25. de Silva R, Cielniak G, Gao J F. Vision based crop row navigation under varying field conditions in arable fields. *Computers and Electronics in Agriculture*, 2024, **217**: 108581
26. Shi J Y, Bai Y H, Zhou J, Zhang B H. Multi-crop navigation line extraction based on improved YOLO-v8 and threshold-DBSCAN under complex agricultural environments. *Agriculture*, 2024, **14**(1): 45
27. Xu Y, Xue X Y, Sun Z, Gu W, Cui L F, Jin Y K, Lan Y B. Global path planning for navigating orchard vehicle based on fruit tree positioning and planting rows detection from UAV imagery. *Computers and Electronics in Agriculture*, 2025, **236**: 110446
28. Lin Y D, Xia S Y, Wang L X, Qiao B Y, Han H, Wang L H, He X K, Liu Y J. Multi-task deep convolutional neural network for weed detection and navigation path extraction. *Computers and Electronics in Agriculture*, 2025, **229**: 109776
29. Li D F, Li B L, Feng H Q, Kang S, Wang J, Wei Z B. Low-altitude remote sensing-based global 3D path planning for

- precision navigation of agriculture vehicles—Beyond crop row detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2024, **210**: 25–38
30. Ding H Z, Zhang B H, Zhou J, Yan Y X, Tian G Z, Gu B X. Recent developments and applications of simultaneous localization and mapping in agriculture. *Journal of Field Robotics*, 2022, **39**(6): 956–983
31. Zhuang L C, Zhong X R, Xu L J, Tian C B, Yu W S. Visual SLAM for unmanned aerial vehicles: localization and perception. *Sensors*, 2024, **24**(10): 2980
32. Liu M R, Zhang W J, Wei C, Bao Z M, Hu J J, Zhou J W. PLGAT: underwater *Plectropomus leopardus* recognition using global attention mechanism and transfer learning. *IEEE Access*, 2024, **12**: 185149–185159
33. Qi Y L, He Y T, Qi X M, Zhang Y, Yang G Y. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In: Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: *IEEE*, 2023, 6047–6056
34. Vidović I, Cupec R, Hocenski Ž. Crop row detection by global energy minimization. *Pattern Recognition*, 2016, **55**: 68–86
35. Su R, Gao L, Zhu Z T, Liu C, Chen D. Canopy-under navigation path recognition method for human-machine collaborative harvesting robots based on improved DeepLabV3+. *Transactions of the Chinese Society for Agricultural Machinery*, 2026, **57**(6): 1–12 (in Chinese)
36. Su T, Wang L, Ban C, Chi R J, Ma Y Q. Interrow path navigation line detection of maize in middle and late period based on semantic segmentation. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, **55**(10): 275–285 (in Chinese)
37. Woo S, Park J, Lee J Y, Kweon I S. CBAM: convolutional block attention module. In: Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: *Springer*, 2018, 3–19
38. Ouyang D L, He S, Zhang G Z, Luo M Z, Guo H Y, Zhan J. Efficient multi-scale attention module with cross-spatial learning. In: Proceedings of 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island, Greece: *IEEE*, 2023, 1–5
39. Dong Z D, Zhu D Y, Huang M, Lin Q F, Møller-Jensen L, Silva E A. MSCANet: multi-scale spatial-channel attention network for urbanization intelligent monitoring. *Remote Sensing*, 2026, **18**(1): 159
40. Han K, Wang Y H, Guo J Y, Wu E H. ParameterNet: parameters are all you need for large-scale visual pretraining of mobile networks. In: Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: *IEEE*, 2024, 15751–15761
41. Yang J N, Liu S L, Wu J J, Su X Y, Hai N, Huang X L. Pinwheel-shaped convolution and scale-based dynamic loss for infrared small target detection. In: Proceedings of the 39th AAAI Conference on Artificial Intelligence. Philadelphia, PA, USA: AAAI, 2025, 9202–9210
42. Chen L, Xue Z S, Liu X F, Chen M J, Huang H. A novel safety helmet wear detection method based on improved RFBnet. *Journal of Physics: Conference Series*, 2020, **1617**(1): 012035
43. Ye L, Li J, Yin L, Gao J F, Tang Y C. KBQ-RRT*: a smoothness-enhanced kinematic bidirectional quick-RRT* via dual-tree optimization for autonomous navigation in complex orchards. *Journal of Field Robotics*, 2026, **43**(3): 1823–1843