Fruit
Research

RESEARCH HIGHLIGHT

Open Access

https://doi.org/10.48130/frures-0026-0024
Fruit Research **2026**, 6: e028

Performance evaluation of a DCD-YOLO model in grape disease detection

Xiangdong Wang[#], Liu Yang[#], Yi Wu[*], Jiaqi Liao, Yupu Xiong, Yuancheng Wen, Jiqing Peng[*], Song Sheng[*], Jianjun Li[*] and Sen Wang[*]

School of Computer and Mathematics, Central South University of Forestry & Technology, Changsha, Hunan 410004, China
[#] Authors contributed equally: Xiangdong Wang, Liu Yang
* Correspondence: wuyi3721@csuft.edu.cn (Wu Y); pengjiqing17@csuft.edu.cn (Peng J); songsheng@csuft.edu.cn (Sheng S); lijianjun_21@163.com (Li J); csuftwangsen@163.com (Wang S)

Abstract

In response to the severe impact of common diseases such as measles, wilt disease, and rotten holes on grape yield and quality in grape cultivation, as well as the limitations of existing detection models in identifying subtle lesions and locating irregular lesions, this study aims to propose an improved YOLOv11n-based detection model for grape leaf diseases. The goal is to enhance the accuracy and robustness of grape leaf disease detection, thereby providing technical support for the early diagnosis, prevention, and control of grape diseases. An improved model was constructed based on the YOLOv11n algorithm, with key improvements including: embedding the DPPA dual-channel attention module in the backbone network to fuse multi-scale features through dual pathways of local detail perception and global context capture, thereby enhancing the discriminative ability for heterogeneous lesions; replacing the original C3K2 structure with the C2f-DCN module, which introduces deformable convolution to adaptively capture the morphology of irregular lesions and improve localization accuracy. A grape leaf disease dataset (GLDDD) containing 5,000 images was used, divided into training, validation, and test sets in an 8:1:1 ratio. Combined with data augmentation techniques such as Mosaic-9 and Randaugment, the model was trained with 250 training iterations. The improvement effects were verified through ablation experiments and comparative experiments. The following were observed: (1) The performance of the improved model was significantly enhanced, with an mAP of 95.7%, showing excellent performance among similar models. (2) Ablation experiments indicated that the C2f-DCN module had a prominent effect in improving precision and mAP, with minimal impact on parameter count and model size; the DPPA module could improve precision to a certain extent, and their combination effectively optimized model performance. However, although the DFF module improved recall, it significantly increased the parameter count and model size while leading to a decrease in precision. (3) Comparative experiments showed that the improved DCD-YOLO model achieved a 4.9% higher detection precision for measles disease than YOLOv11n and a 5.4% higher precision than YOLOv11n-DPPA. The average detection accuracy for the three diseases was significantly higher than that of other comparative models. (4) The model achieved the optimal performance balance at 250 training epochs, with a precision of 92.3%, a recall of 90.1%, and an mAP of 95.7%, demonstrating the best performance in detection accuracy and coverage capability. The improved YOLOv11n model proposed in this study effectively addresses the issues of difficulty in identifying subtle grape leaf lesions and low localization accuracy of irregular lesions by introducing the DPPA and C2f-DCN modules. It significantly enhances the performance of grape leaf disease detection, meets the practical needs of grape disease detection, and provides a feasible technical solution for the automated detection of crop diseases in the agricultural field.

Keywords: Deep learning, YOLO, Grape, Disease detection, DPPA dual-channel attention mechanism

Citation: Wang X, Yang L, Wu Y, Liao J, Xiong Y, et al. 2026. Performance evaluation of a DCD-YOLO model in grape disease detection. *Fruit Research* 6: e028 https://doi.org/10.48130/frures-0026-0024

Introduction

Grapevine leaf diseases such as rotten holes, grapevine gray mold disease, spot diseases, and wilting remain a significant challenge to global viticulture due to their adverse impacts on photosynthesis and overall yield. Numerous studies have explored automatic, early, and precise recognition technologies for these pathologies[1,2]. Techniques based on image processing and machine learning have demonstrated considerable success. For instance, methods combining Wiener filtering, morphological segmentation, and BP neural networks have achieved high classification accuracy in detecting anthracnose and other grape leaf diseases[3]. Convolutional Neural Networks (CNNs), particularly customized architectures like the UnitedModel and Enhanced VGG16, have been reported to yield superior results in detecting grape leaf diseases with accuracies exceeding 98%[4]. These tools not only recognize multiple disease types but also enable real-time application through embedded systems and UAV-assisted precision spraying mechanisms[5,6]. Further, deep learning architectures like DenseNet, Inception-based CNNs, and YOLOv5-CA integrate attention mechanisms or dense connectivity strategies to enhance robustness against environmental variability and have achieved detection precision above 95% in complex vineyard settings[7,8]. Notably, instance segmentation approaches, such as Mask R-CNN, allow pixel-level classification and have been supported by curated datasets featuring detailed annotations across symptomatic grape organs[9]. These models have been instrumental in advancing high-throughput, non-invasive diagnostic systems.

Compared to other crops, such as ginger or general fruit-bearing plants, similar CNN and hybrid models have demonstrated reliable accuracy levels above 90% in classifying diseases like anthracnose and bacterial wilt, suggesting a broader applicability of these frameworks across plant species[10]. Additionally, fuzzy rule-based systems have also been investigated for their potential in handling imprecise symptom descriptions in grape anthracnose diagnosis[11,12]. Importantly, epidemiological studies have quantitatively characterized the development of anthracnose, linking disease progression to temperature, leaf age, and moisture, offering insights into optimal intervention timings[13,14]. Confirmatory field reports continue to document new strains of anthracnose-causing pathogens like *Elsinoë ampelina* and *Colletotrichum gloeosporioides*, reinforcing the need for scalable, adaptable detection models[15,16].

The conventional reliance on expert-based visual inspection for crop disease diagnosis has long been hampered by subjectivity, labor intensity, and the inability to scale for real-time monitoring. The advent of deep learning, particularly the YOLO (You Only Look Once) object detection framework, has rapidly advanced automated plant disease recognition, offering a balanced trade-off between accuracy and computational speed. Numerous studies have validated the efficacy of various YOLO versions across crops and disease types. For instance, YOLOv5 was reported to achieve early-stage plant disease detection at 45 FPS, enabling rapid field deployment with promising accuracy and cost-effectiveness[16]. Comparative evaluations between YOLOv5, v6, v8, and v9 demonstrated that YOLOv8s and YOLOv9-C delivered high precision and recall on diverse plant disease datasets such as PlantDoc, indicating strong generalizability[17]. In citrus crops, YOLOv8 outperformed older versions by achieving over 96% mean average precision across several disease classes[18], while its deployment on low-power platforms like Raspberry Pi still maintained high detection efficiency. In addition to performance benchmarks, architectural enhancements have been introduced to adapt YOLO to the complexities of agricultural conditions. For example, SE-YOLO, a lightweight adaptation of YOLOv8, integrated GSConv and multi-scale attention modules to significantly improve accuracy in small-object rice disease detection while reducing computational load[19]. YOLO-Pest, built on YOLOv4, improved multi-class pest detection by incorporating MobileNetv3 and feature pyramid networks to address scale variance challenges in field imagery. Meanwhile, industrial-scale adaptations like YOLOv5-CAcT introduced model compression and quantization for real-time deployment with only a 2 MB model size and 1.563 ms inference time, achieving 94.24% accuracy across 59 crop disease categories[20].

Transfer learning further enhanced YOLO-based models; fine-tuning YOLOv5 variants resulted in F1 scores exceeding 0.96 and detection accuracies above 98% on plant disease datasets, with clear implications for deployment in UAVs and mobile platforms[21]. Cross-crop studies reveal the wide applicability of YOLO-based detectors: in cassava, YOLOv9 variants surpassed 80% mAP in detecting multiple disease classes[22], while YOLO-NAS achieved state-of-the-art precision and F1 scores in fava bean rust identification, highlighting YOLO utility beyond viticulture. Further reviews consolidate YOLO position as the most promising one-stage detector in precision agriculture, with practical successes across crops including rice, soybean, and grapevine, albeit challenges in symptom overlap and data scarcity remain[23].

Despite the advantages of YOLO-based models in plant disease detection, their performance in real-world agricultural environments remains constrained by several persistent challenges, notably the detection of early-stage, small, and irregularly shaped lesions; the impact of complex lighting conditions; and the morphological similarity between different disease types and healthy leaf textures. These limitations have been repeatedly observed in diverse applications. For instance, while YOLO has proven effective in high-frame-rate detection, its performance drops significantly in complex field settings where lighting and background noise introduce high visual variability[16]. Enhanced architectures have been developed to address these issues, such as CEFW-YOLO, which introduced attention mechanisms and fine-grained feature extraction to handle diverse leaf morphologies and illumination variance, improving mAP by over 7% compared to baseline YOLO11n. Studies have also shown that early-stage lesion detection remains problematic due to the small size and ambiguous visual characteristics of disease symptoms. In response, model adaptations incorporating channel and spatial attention, such as those applied in YOLOv8-ASFF for tea diseases, significantly enhanced detection under low-light and small-object conditions, yielding mAP values over 95% and outperforming other models by up to nine percentage points[19]. Similarly, optimized YOLOv8 heads have been proposed to address object scale disparity and occlusion in strawberries, resulting in improved precision for small-object detection without sacrificing accuracy on larger targets. The intrinsic visual similarity between different diseases and healthy tissues has been shown to confound YOLO-based detectors. To mitigate this, hybrid models like HODRF have been developed, combining YOLOv7 with EfficientNetV2 in a hierarchical detection-classification structure, achieving up to 7.5 points improvement in F1 scores over single-stage detectors by reducing false positives on healthy samples[24]. More generally, systematic reviews emphasize that while YOLO excels in controlled environments, real-world applications are hindered by limited annotation quality, data imbalance, and feature overlap between disease stages and classes, necessitating the integration of multimodal inputs and adaptive learning mechanisms[23]. Improved versions like YOLOv8 and YOLOv9 have made strides in reducing false detections and boosting robustness through enhanced feature fusion and attention layers[17]. Ultimately, while recent advancements in YOLO variants and attention-augmented models have notably improved performance in real-world scenarios, challenges relating to lesion complexity, environmental noise, and class ambiguity remain active research frontiers in automated plant disease detection.

In summary, this study proposes an enhanced YOLOv11n-based detection architecture, DCD-YOLO, specifically optimized for grape leaf disease recognition in complex field environments. By integrating the Dual-Path Perception Attention (DPPA) mechanism and the C2f-DCN deformable convolution module, the model effectively addresses the intrinsic challenges posed by grape leaf lesions, such as small early-stage disease features, irregular morphological patterns, and environmental interferences like strong light and shadow occlusion. Comprehensive evaluation across multiple metrics demonstrated that the synergistic interaction of DPPA and C2f-DCN not only enhanced the model's precision in capturing fine-grained lesion textures and complex lesion morphologies but also significantly improved its robustness against background noise and occlusion. The improved model achieved an mAP of 95.7%—a marked improvement over baseline YOLOv11n and YOLOv11n-DPPA models—and showed superior lesion localization fidelity, particularly under strong light and partial occlusion, as evidenced by heatmap and feature map visualizations. Confusion matrix analysis confirmed its accurate classification across anthracnose, measles, and wilt with minimal misclassification, while PR curve and ablation experiments further verified the model optimization in balancing detection accuracy and inference efficiency[25]. Though the DCD-YOLO model demonstrated a trade-off in inference speed compared to lighter architectures, its performance at 250 training epochs presented an optimal balance between accuracy, robustness, and generalization. These findings establish a solid foundation for deploying high-precision, lightweight, and resilient plant disease detection models in real-time agricultural applications, and lay the groundwork for future enhancements involving model compression, multimodal sensor integration, and disease grading for intelligent decision support[26].

Materials and methods

Data acquisition

The dataset used in this study was collected from the CSDN website platform. The CSDN public database contains grape leaf disease images covering different plants, various disease types, and diverse scenarios. The data path is https://download.csdn.net/download/m0_64879847/88535108. In this study, leaf images of three frequently occurring and loss-prone diseases in grape diseases, namely measles[27,28], wilt[29], and rotten holes[14,30], as well as self-collected images of healthy leaves, were gathered from the CSDN public database. The images of each category are shown in Fig. 1.

After downloading public dataset images from the CSDN and real disease images obtained through web crawlers, the images were screened, and a total of 5,000 images remained after blurry photos were removed to construct a dataset for grape leaf disease detection (grape leaf disease detection dataset) GLDDD.

Model construction and verification

The DPPA attention module is embedded in the backbone network, and multi-scale features are fused through dual paths (local detail perception path + global context path), which significantly enhances the ability to distinguish heterogeneous lesions such as measles disease, diffuse spots, and rotten-hole perforations. The C2f-DCN module is used to replace the original C3K2 structure, and deformable convolution is introduced to adaptively capture the irregular morphological features of lesions, thereby improving the positioning accuracy of diverse lesions in wilt disease; the final size of the model is 28.36 MB. The GLDDD dataset (5,000 images) is divided into a training set, a validation set, and a test set in an 8:1:1 ratio to ensure class balance (stem pock disease, wilt disease, rotten holes). Among them, the training set contains 4,000 images, the validation set has 500 images, and the test set consists of 500 images. The detailed division is shown in Table 1.

Mosaic9 data augmentation (random scaling, rotation, HSV color perturbation) was used to simulate complex field environments and enhance generalization. Image size was unified to 640 × 640 pixels; the number of model training iterations was set to 250; 40% random erasure was used to simulate occlusion and enhance the model's anti-interference ability; Randaugment was used to automatically select the appropriate combination of augmentation operations. The final model mAP values were 95.7%, indicating high precision.

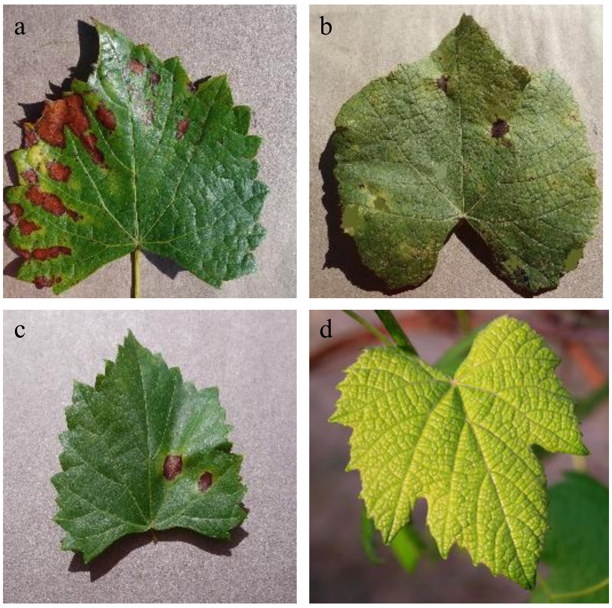

Fig. 1 Display of three diseased and healthy leaves. (a) Measles. (b) Wilt. (c) Rotten holes. (d) Healthy.

Improved detection of grape leaf diseases by YOLOv11n

Introduction to the YOLOv11n algorithm model

YOLOv11, a new-generation object detection algorithm with multi-scale versions, inherits the YOLO series' efficient detection advantages and has optimized network structure and performance, showing good applicability in agricultural disease detection like grape leaf disease recognition, but faces challenges such as low recognition accuracy for subtle lesions, insufficient positioning accuracy for diverse morphological lesions, weak generalization ability for minority-class diseases under imbalanced samples, susceptibility to background noise and lighting changes, and potential real-time performance issues on resource-constrained devices, requiring fine-tuning strategies to enhance feature capture and detection robustness.

The improved network

To enhance the accuracy of grape leaf disease detection, this study proposes an improved and optimized network architecture. The model is mainly based on the YOLOv11n algorithm and further improved and optimized on it, as shown in Fig. 2.

In the YOLO11n network architecture, the C2f-DCN module is embedded into the core feature extraction layers of the backbone: the original feature fusion modules are replaced at the 4th layer (P3/8 scale with parameters [512, False, 0.25]), the 6th layer (P4/16 scale with parameters [512, True]) and the 8th layer (P5/32 scale with parameters [1024, True]), respectively. The feature alignment capability is enhanced by introducing deformable convolution. The DPPA module is configured as [256, 256] (both input and output channels are 256) and accurately inserted into the 11th layer of the backbone (after the SPPF and C2PSA layers), which undertakes 1,024-dimensional features and completes dimension adaptation and feature enhancement. The overall architecture retains the original P3-P5 multi-scale output structure of YOLO11, and the insertion of C2f-DCN and DPPA does not alter the feature flow from the backbone to the head. The head layer still accomplishes multi-scale feature fusion through upsampling, feature concatenation, and the C3k2 module, and the Detect layer finally outputs detection results at three scales of P3/8, P4/16, and P5/32.

Specifically, the following improvements were made compared to the original model:

(1) Backbone: A DPPA dual-channel attention mechanism module is added to the feature fusion section of the backbone. It fuses detail and context paths to generate a spatial attention weight map for

Table 1. Dataset of grape leaf diseases.

Disease category	Train set	Valid set	Test set	Combined
Measles	1,333	167	167	1,667
Wilt	1,333	167	167	1,667
Rotten holes	1,334	166	166	1,666
Total	4,000	500	500	5,000

https://doi.org/10.48130/frures-0026-0024

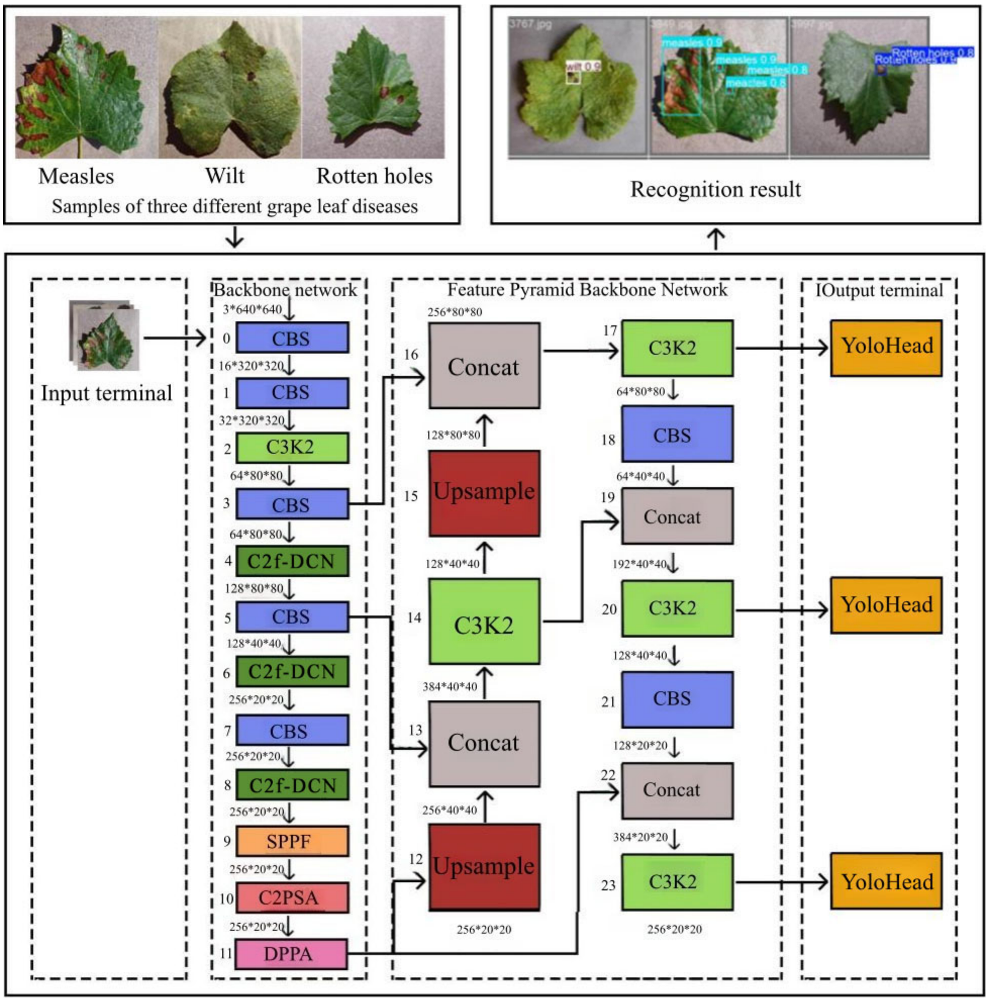

Fig. 2 Flowchart of the improved YOLOv11n algorithm.

feature enhancement, while retaining the SPPF module to solve the problem of repetitive feature extraction, improving the speed of candidate box generation and reducing computational costs.

(2) C2f-based part: While retaining the efficient feature fusion and multi-layer Bottleneck feature transfer enhancement capabilities of the C2f module, deformable convolution is introduced to replace the C3K2 modules in the 2nd, 3rd, and 4th layers of the backbone network. It adaptively adjusts the sampling position of convolution kernels by learning offsets, accurately captures the irregular shapes of grape leaf diseases (such as rotten holes perforation and irregular wilting of wilt), improving spot localization accuracy and the model robustness in capturing complex disease features.

Dual-channel attention mechanism (DPPA)

DPPA (Dual Path Attention) is an attention module designed for the complexity of grape leaf disease characteristics, achieving precise focus and multi-scale information fusion of lesion features through a parallel 'local detail perception' and 'global context capture' dual-path structure, significantly enhancing the model's ability to discriminate heterogeneous lesions. The core design and working mechanism are as follows:

(1) Module structure design: The DPPA module, centered on dual-channel parallel processing, comprises a feature input layer, dual-path extraction unit (local detail perception via small receptive field convolution with multi-level residual connections, global context capture via large receptive field convolution + Global Average Pooling), feature fusion layer, and attention weight generation layer, which concatenates/aggregates dual-path features to generate multi-scale fused maps and applies sigmoid activation for spatial attention weight maps distinguishing key lesion areas.

(2) How it works: The DPPA module operates through three steps: dual-path extraction of lesion fine-grained and global associative features, dynamic weight learning (assigning high weights to key lesion areas and low weights to background/non-critical regions) via activation function-generated weight maps, and pixel-wise multiplication of the original feature map with the weight map to enhance key lesion features and suppress background noise before outputting the enhanced features.

(3) Core features and advantages: The DPPA module enhances fine spot recognition (increasing measles detection accuracy by 0.9 percentage points) and global discrimination of complex diseases in grape leaf disease detection, while featuring a lightweight design (2.09 M parameters, 11.8% of DFF modules; 4.41 MB model size increase) to balance performance and efficiency for field mobile device applications.

C2f-DCN module optimization

To improve the model localization accuracy and feature capture ability for various morphological lesions in grape leaf diseases (such as perforation in rotten holes and irregular wilting in fusarium wilt), this study proposes the C2f-DCN module by introducing Deformable Convolution on the basis of the C2f module. Instead of the traditional convolution in the original C3K2 module, adaptive extraction of irregular lesion features is achieved. Its design principles, structural improvements, and functional advantages are shown in Fig. 3.

As shown in the figure, the design principles and functional modules can be deeply analyzed from three aspects: process logic, core components, and improvement logic:

(1) Process logic: the C2f-DCN module follows the process of 'input preprocessing → dual-path feature enhancement → feature aggregation output' to meet the requirement of 'multi-morphology lesion feature extraction' in grape leaf disease detection. The specific steps are as follows:

① Input preprocessing (C2f-DCN → CBS → Split) C2f -DCN: Receive the feature map output from the upstream network (such as the lesion features at different levels in the backbone) as the initial input for module processing.

② CBS (Convolution + Normalization + Activation): Perform basic convolution (typically with BN layers and activation functions) on the input feature map, compress the number of channels, and initially extract the basic features of the lesion (such as color, texture) to provide a more concise input for subsequent deformable convolution (DCN).

③ Split (Feature Splitting): Split the feature map into 'direct transfer branches' and 'DCN-enhanced branches'. Direct transfer branches: retain the basic information of the original features (such as the texture of healthy leaves, the background of disease spots), and avoid the loss of details caused by deep convolution (suitable for the complex scenario of 'mixed disease spots and healthy areas of grape leaves'). Dcn-enhanced branching: enter the Bottleneck-DCN unit and focus on the irregular morphological features of the lesion (such as rotten holes, perforation edges, wilted curled contours) through deformable convolution (DCN).

(2) Core component analysis: The Bottleneck-DCN unit consists of *n* Bottleneck-DCN modules in series, each containing double CBS, DCN, and residual connections: double CBS extracts common lesion features, DCN dynamically adjusts convolution kernel sampling coordinates via learned offsets to fit irregular grape leaf lesion shapes (e.g., serrated edges of rotten holes perforations), and residual connection retains basic features while enhancing morphological information; then, Concat fuses the direct transfer branch and DCN-enhanced branch at the channel dimension, and the output CBS compresses channels to provide lightweight and informative features for downstream networks.

(3) Combining the pain points of 'morphological diversity and difficulty in distinguishing from healthy areas' of grape leaf diseases (rotten holes, wilt, measles, etc.). The design advantages of the C2f-DCN module can be summarized as:

① Accurately capture the morphology of the lesion. The fixed convolution of the traditional C2f module cannot fit the irregular edges of the lesion (such as the irregular contours of the perforation of rotten holes). DCN uses dynamic offset sampling to make the convolution kernel actively 'fit' the morphology of the lesion, solving the problem of feature loss caused by 'mismatch between the sampling position and the lesion'.

② Balancing feature completeness with lightness, through 'feature splitting + dual-branch fusion', it retains the basic information of healthy leaves and the background of the lesion (direct transfer branch) while enhancing the morphological details of the lesion (DCN branch), avoiding ignoring background associations due to excessive focus on the lesion (such as the auxiliary role of the spatial relationship between the lesion and the leaf vein in recognition). Residual connections and channel compression (CBS) control parameter growth (lighter than pure DCN modules) to meet the real-time requirements of field detection equipment.

Model training and evaluation metrics

Training environment and parameter settings

The model in this study was trained on the following hardware settings: processor 11th Gen Intel(R) Core(TM)i5-11260H@2.60 GHz (2.61 GHz), with 16.0 GB of memory, and operating system 64-bit (x64-based processor). The GPU is an RTX 3050, the NVIDIA graphics driver is v512.15, and CUDA is v11.5.

The GLDDD dataset (8:1:1 split into 4,000/500/500 samples) was resized to 640 × 640. Mosaic9, random scaling/rotation/HSV perturbation, and 40% random erasure were used for augmentation. Training: 250 epochs, initial LR 0.01 (SGD, cosine annealing), batch size 16.

Results and analysis

Confusion matrix and heat map analysis

This study is based on the improved YOLOv11n with the addition of DPPA and the replacement of the C2f-DCN module, using the standard model parameter sizes, as shown in Fig. 4, which is the normalized confusion matrix diagram of the improved model standard parameters.

The model achieves recognition accuracies of 95%, 89%, and 98% for rotten holes, measles, and wilt disease respectively, with 24%, 44%, and 32% of the three types of diseases misclassified as background, and 5%, 11%, and 2% of background misjudged as the corresponding diseases, indicating that the model has varying degrees of confusion in distinguishing diseases from the background.

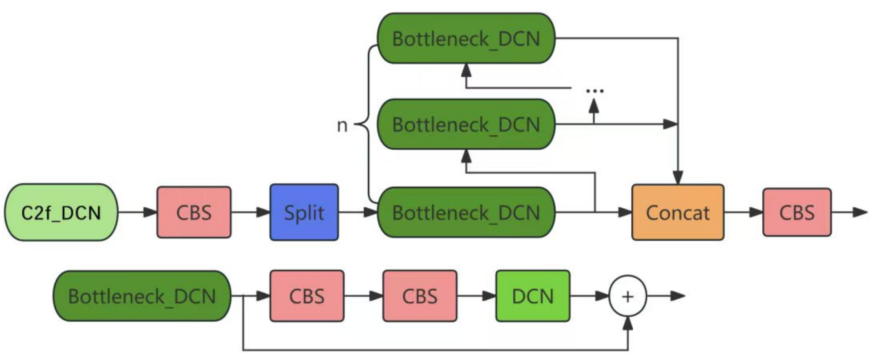

Fig. 3 C2f-DCN module structure diagram.

https://doi.org/10.48130/frures-0026-0024

Heat map analysis of different models

In this study, the heat map is used to visually reflect the degree of attention the model pays to the image area. Dark areas represent the key feature areas of the lesion determined by the model, and light or colorless areas represent the background or secondary feature areas. As shown in Fig. 5, the comparison of heat maps of different models in the three scenarios of the original image, strong light exposure, and shadow occlusion can be used to analyze the localization accuracy and anti-interference ability of each model for the lesion, as follows:

Without environmental interference, DCD-YOLOs—the optimized model—achieves the most accurate lesion focus in heatmaps and excels at parsing complex lesion features.

Under strong light, YOLOv11n and YOLOv11n-DFF are susceptible to interference, while the improved DCD-YOLO and DCD-YOLOs have strong anti-interference capabilities, with DCD-YOLOs reaching the highest confidence of 0.85.

In shadowed backgrounds, YOLOv11n and YOLOv11n-DFF have high missed detection rates; the improved models perform outstandingly, with DCD-YOLOs achieving the highest confidence of 0.86 for obscured lesions and the lowest missed detection rate, making the improved YOLOv11n models significantly superior overall.

Visual analysis of feature maps

Feature map visualization can visually reflect the ability of the model to capture and process disease features at different network levels. By observing the response intensity in the feature map, the extraction logic of the model for key information of the lesion can be analyzed. Shown in Fig. 6 are the 16 channel images of the disease photo feature map from the shallow to the deep layers of the four network layers in model recognition.

In this study, the network progressively extracts features of grape rotten-hole leaves from shallow to deep layers: the shallow network (Fig. 6b) extracts basic visual features such as edge contours, generating continuous dark responses; the middle-layer network (Fig. 6c) fuses basic features and leverages the local detail perception capability of the DPPA module to enhance responses to fine rotting textures at perforation edges, enabling the distinction between the 'lesion body' and the 'cavity area'; the deep network (Fig. 6d, e) further extracts the semantic feature of 'perforation + rot' as well as high-order semantic features. It reflects the learning of the typical rotten-hole feature of 'coexistence of multiple perforations' through 'regional aggregation', and can produce differentiated responses based on lesion severity, providing crucial support for disease identification and differentiation.

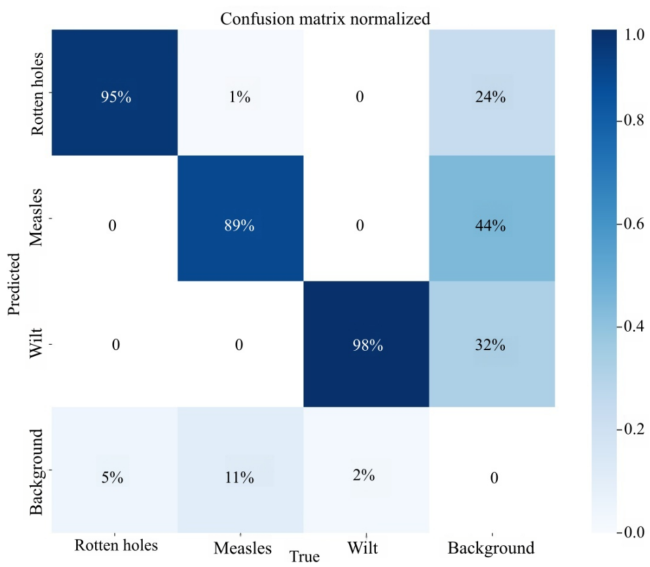

Fig. 4 Normalized confusion matrix diagram.

Ablation test

To verify the effectiveness of the improved YOLOv11n, ablation tests were conducted on each of its improved modules in this study, and the results are shown in Table 2.

Overall, the C2f-DCN module excelled in improving model accuracy and mAP, with little impact on the number of parameters and model size. The DPPA module can also improve accuracy to some extent. Although the DFF module can improve recall, it will significantly increase the number of parameters and model size, and lead to a decrease in accuracy. Adding both DPPA and C2f-DCN modules can effectively enhance model performance, and YOLOv11s further optimizes on this basis to achieve better detection results.

Comparative test analysis of different models

To further validate the metrics and performance of the improved YOLOv11n algorithm model in this study, the improved model and its original model YOLOv11n, the model YOLOv11n-DPPA with only DPPA (dual-channel attention mechanism) added, and the DCD-YOLO module with DPPA added under default model parameters, and with the 2nd, 3rd, and 4th C3K2s in the original model replaced with the C2f-DCN module. The models were compared, trained, and tested on the training and test sets of the Grape Leaf Spot Detection Dataset (GLDDD) with the same number of training rounds and default parameters. The following figures show the comparison of test results for fusarium wilt, measles, and rotten holes on the YOLOv11n, YOLOv11n-DPPA, DCD-YOLO, and the improved DCD-YOLO models, respectively. The test results are shown in Fig. 7.

The study effectively enhanced the ability to extract lesion features from grape leaves and improved issues such as false detections and missed detections by fusing the dual-channel attention mechanism DPPA with the C2f-DCN module (where the C3K2 modules in the 2nd, 3rd, and 4th layers of the backbone network were replaced). Specifically, YOLOv11n-DPPA, which incorporates DPPA, showed a significant improvement in detection accuracy compared to the original YOLOv11n; the further improved DCD-YOLO achieved higher accuracy and stronger robustness than YOLOv11n-DPPA. These results confirm that the optimization strategies can significantly improve the detection accuracy of the model.

Table 3 shows the comparison of the average accuracy detection effects of different algorithm models on the test set, and tests the actual detection performance of different models.

It can be seen from the table that the average detection accuracy of the improved YOLOv11s model for different diseases has reached 95.7%, which is a significant improvement compared to other algorithm models.

Comparative test analysis of the training performance of different models

To compare the actual values of precision and recall of the three types of grape leaf disease detection targets, the R and P curves of these targets were visualized as shown in Fig. 8.

It can be seen from the figure that the recall and precision of measles are low because the lesions themselves are fine and similar

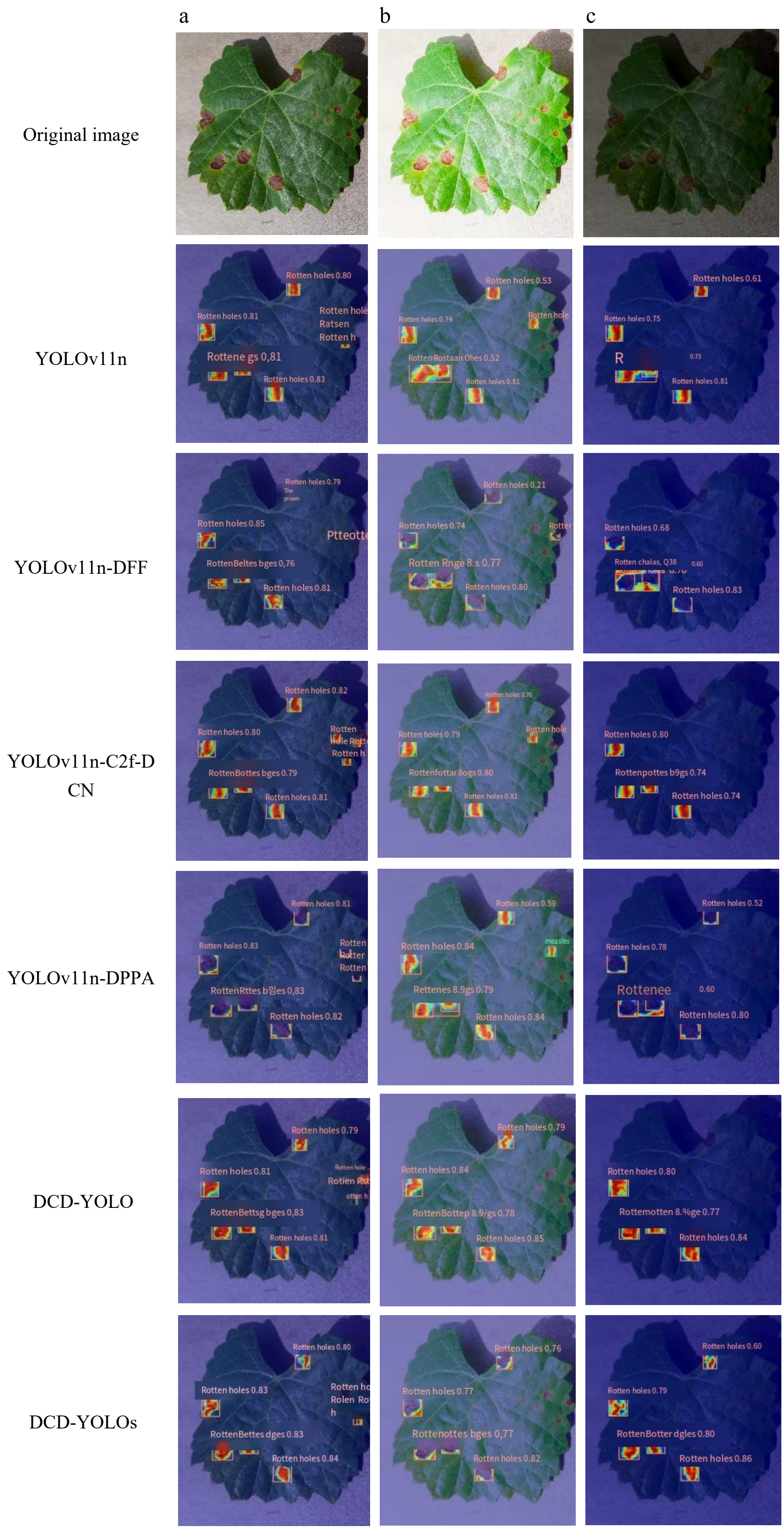

Fig. 5 Different models for disease detection under different background conditions. (a) Original image. (b) Against a strong light background. (c) Against a shaded background.

https://doi.org/10.48130/frures-0026-0024

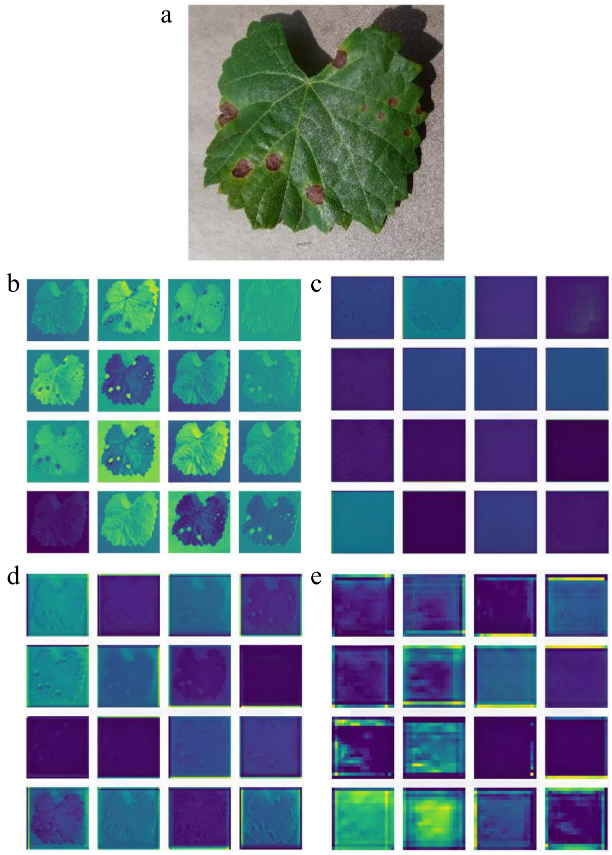

Fig. 6 Feature map visualization.

to the texture of health, combined with external environmental interference, and the visual features of measles lesions are easily disturbed by lighting conditions (such as strong light causing the color of the rash to become lighter, weak light causing the contrast to decrease) and background noise (such as dust and shadows on the leaves), which mask the characteristics of the lesions.

We compared the PR curve diagrams for detecting three diseases on grape leaves using different algorithms to directly reflect the overall detection accuracy of the model through the size of the area enclosed by the curves and the coordinate axes. As shown in Fig. 9, it is used to quantify the overall accuracy of the model, compare the recognition performance of different types of diseases, and provide an intuitive basis for model optimization.

Taking recall as the horizontal coordinate and precision as the vertical coordinate, the PR curve comparison shows that the improved algorithm in this paper, especially the DCD-YOLO, has a significantly larger area enclosed by the PR curve with the coordinate axis, and its accuracy in identifying grape leaf measles spots is improved by 4.9% compared to YOLOv11n and 5.4% compared to YOLOv11n-DPPA, demonstrating better model accuracy.

As shown in Fig. 10, the four algorithms were trained for 250 rounds. The training process for disease detection generates the training set loss curve, the mean curve of the training set object detection loss, the training set loss curve, the precision curve, the recall curve, the validation set loss curve, the mean curve of the validation set object detection loss, the mean curve of the validation set classification loss, the mAP curve, and different types of mAP curves such as the 0 curve.

It can be seen from these curves that the loss curve of the improved YOLOv11s object detection algorithm used in this paper drops significantly faster, and the area enclosed by mAP is significantly better than that of the other object detection algorithms.

To verify the usability of the improved YOLOv11n algorithm, the DCD-YOLO algorithm model was trained and analyzed with a gradient of 50 rounds, as shown in Table 4, for performance comparison at different rounds.

Comparing the overall performance of each round in the table, the model performed significantly better at 250 rounds than other rounds, specifically as shown in:

① The highest core accuracy index: mAP reached 95.3%, the maximum of all rounds in the table, and was very close to 95.7% of the improved YOLOv11s, indicating that the model had the best classification accuracy for diseases at this time.

② The balance of precision and recall is optimal: Precision (92.3%) and recall (90.1%) are both second-highest in the table (only 300 rounds recall is 0.5 percentage points higher, but precision and mAP are lower), with a small gap (2.2 percentage points), indicating that the model achieves the best balance between 'reducing misjudgments' and 'reducing missed detections'. It meets the core demand for 'precise spot identification' in grape leaf disease detection.

To verify the superiority of the improved YOLOv11s algorithm, the curves of the above algorithm model after 250 rounds of training were collated and analyzed, as shown in Table 5, which is a performance comparison table of the three algorithms.

The average precision and average recall of the improved YOLOv11s for grape leaf disease detection were 92.0% and 91.9%, respectively, and the mAP precision of the improved YOLOv11s was 95.7%. The improved YOLOv11s model was more accurate than the rest of the algorithm models.

By testing and summarizing the performance of various models, the average accuracy of leaf disease detection in the improved YOLOv11s was increased by 1.60, 1.50, and 1.20 percentage points, respectively, compared to the YOLOv11n-DFF, YOLOv11n-DPPA, and YOLOv11n models, reaching 95.7%. The frame rate was also improved.

Discussion

In recent studies focusing on precision agriculture, numerous efforts have been made to enhance the robustness and accuracy of

Table 2. Results of ablation trials.

DPPA	DFF	C2f-DCN	*P* (%)	*R* (%)	mAP (%)	Number of parameters (M)	Model size (MB)	Inference speed (ms)	Frames per second (fps)
—	—	—	89.6	90.9	94.5	2.58	5.23	3.5	285.71
√	—	—	90.5	90.2	94.2	4.65	9.64	4.8	208.33
—	√	—	86.2	92.1	94.1	17.67	35.54	11.6	90.91
—	—	√	91.6	90.7	95.0	2.77	5.57	3.4	294.12
√	—	√	92.3	90.1	95.3	5.07	9.98	4.8	208.33
YOLOv11s									
√	—	√	92.0	91.9	95.7	14.68	28.36	9.8	128.21

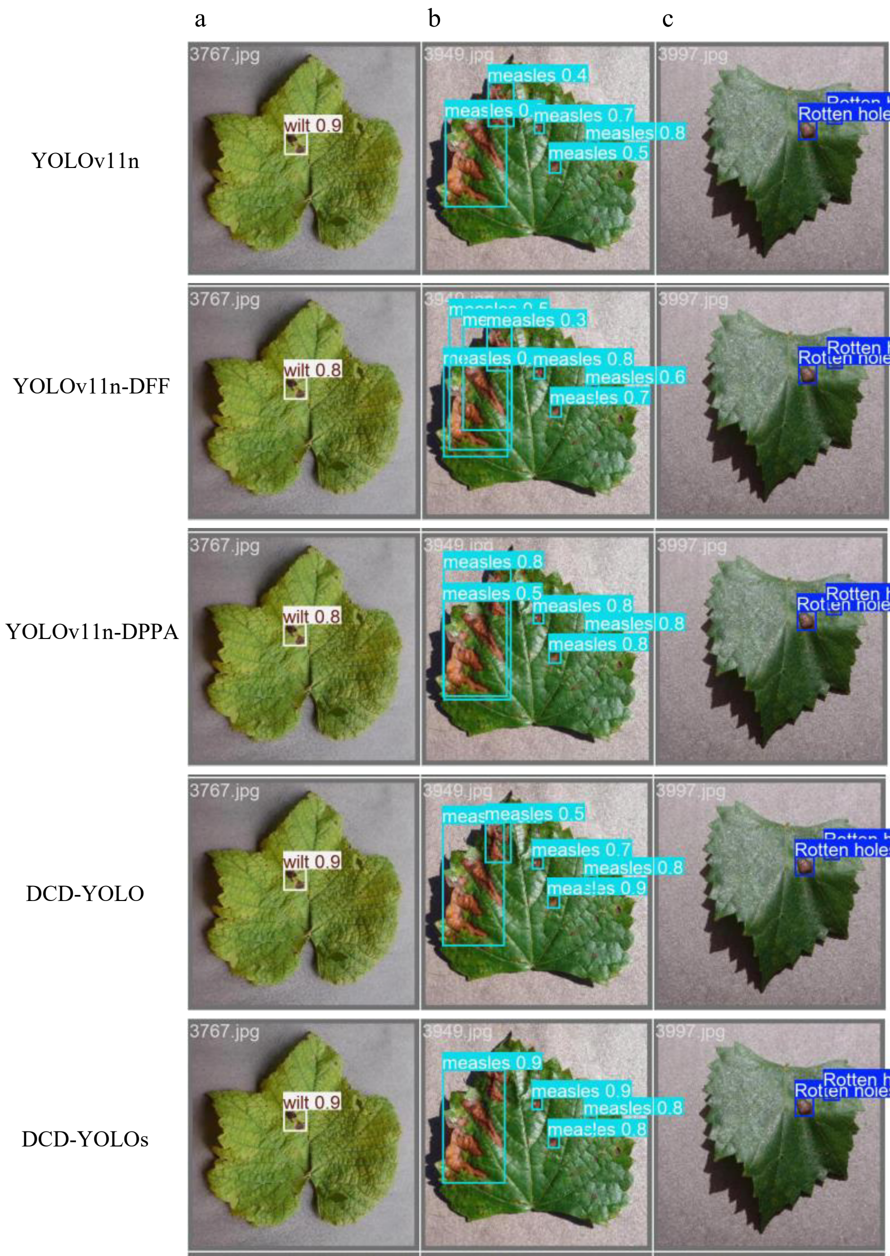

Fig. 7 Model detection of grape leaf diseases (a) Fusarium wilt, (b) measles, and (c) rotten holes.

Table 3. Comparison of average accuracy detection results of different networks on the test set.

Model	Wilt (%)	Measles (%)	Rotten holes (%)	mAP (%)	Inference speed (ms)	Frames per second (fps)
YOLOv11n	0.981	0.870	0.985	0.945	3.5	285.71
YOLOv11n-DPPA	0.978	0.865	0.983	0.942	4.8	208.33
DCD-YOLO	0.983	0.899	0.976	0.953	4.8	208.33
DCD-YOLOs	0.970	0.919	0.982	0.957	9.8	128.21
YOLOv11n-DFF	0.968	0.876	0.978	0.941	11.6	90.91
YOLOv8s	0.981	0.906	0.984	0.957	6.7	149.25
YOLOv8l	0.986	0.883	0.976	0.948	25.1	39.84
YOLOv12	0.977	0.775	0.978	0.910	5.9	169.50

YOLO-based models for plant disease detection through architectural improvements and the integration of attention mechanisms. The newly developed DCD-YOLO model, designed for grape leaf disease detection under complex environmental conditions, aligns with these innovations. By incorporating a dual-path perception attention mechanism and deformable convolution-based modules, this model achieved an outstanding 95.7% mAP, validating the significant gains from structural and attentional enhancements. Similar advancements have been reported in YOLO-GrapeNet, which introduced a Multi-Scale Dilated Attention mechanism and

https://doi.org/10.48130/frures-0026-0024

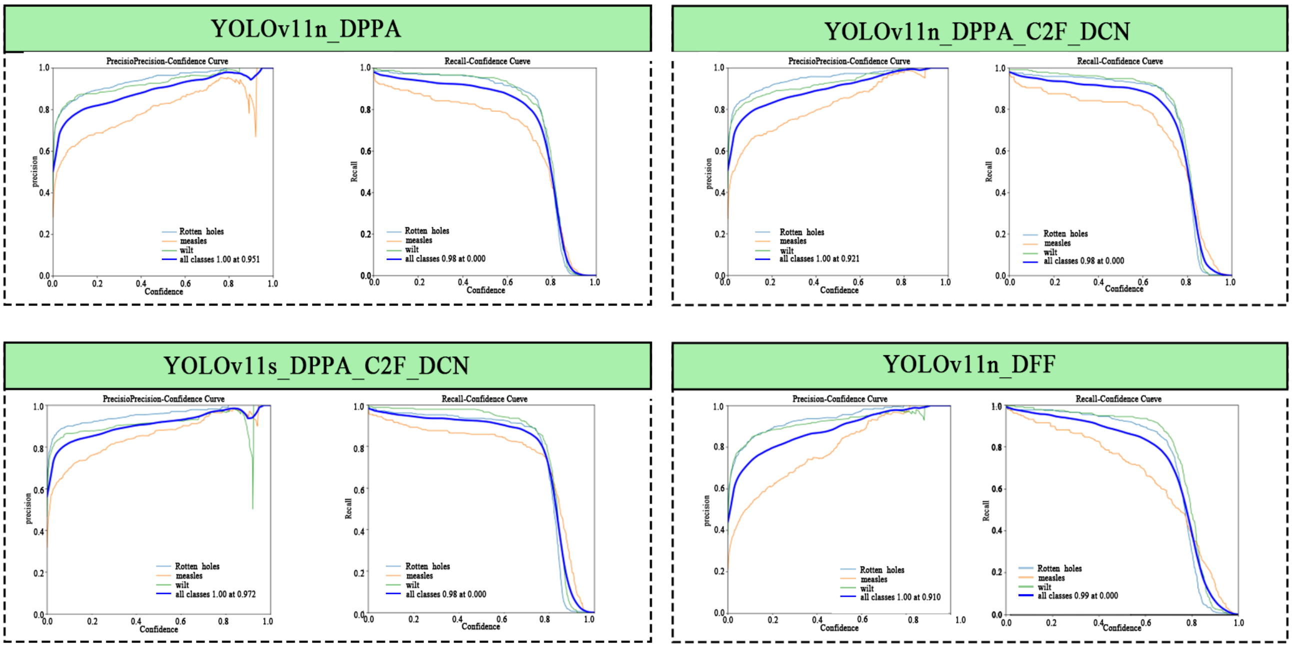

Fig. 8 P and R curves of three types of object detection algorithms.

DilateFormer to extract subtle grape disease features, yielding a mAP of 93.9% and a precision of 92.3%[31]. Other models such as MSAM-YOLO and GCS-YOLO have further explored lightweight convolutional attention modules and Ghost Modules, respectively, to emphasize diseased regions while minimizing model complexity, reaching detection accuracies up to 96.2%[32,33]. The Convolutional Block Attention Module (CBAM) has emerged as a common tool across multiple studies to refine feature extraction, including in GFCD-YOLOXS, which reached a top-tier accuracy of 99.10% in grape leaf classification by enhancing feature fusion and edge detection capability[34]. In models like MSCI-YOLOv8s and the improved YOLOv8 for black rot detection, attention mechanisms like CBAM and Efficient Multi-Scale Attention (EMA) improved the sensitivity to subtle disease markers, increasing detection performance up to 97.7% in real-time applications[35]. Comparable results were also demonstrated in CD-YOLO for tomato disease detection and YOLO-Leaf for apple leaf diseases, underscoring the generalizability of such mechanisms across plant species[36]. Additionally, the incorporation of deformable convolutions and dual-path attention, as seen in the DCD-YOLOs, aligns with improvements in spatial representation and adaptive learning, critical for distinguishing irregular lesion patterns in cluttered environments, reaffirming that attention-guided YOLO architectures represent a pivotal advancement in plant disease diagnostics.

The integration of deformable convolutional networks (DCNs) into YOLO-based plant disease detection models has demonstrated significant effectiveness, particularly in enhancing the model's capacity to recognize irregular lesion morphology and to infer occluded features, both crucial for accurate detection under complex field conditions. The C2f-DCN module in our DCD-YOLOs model exemplifies this by producing dense, contiguous activation maps around irregular rotten-hole lesions and effectively mitigating occlusion-related misdetections. This finding is consistent with the BED-YOLO model for tomato disease detection, where DCNs enhanced the model's adaptability to leaf overlaps and blurred lesion boundaries, yielding a 91.3% mAP and significantly outperforming its baseline in real-world imagery[37]. Similarly, models such as YOLO-DC and ASD-YOLO have leveraged deformable convolutions to adaptively reshape receptive fields, improving accuracy in detecting targets with diverse shapes and occlusions, with reported mAP gains of 3.5%–5.7% across applications[30]. In grape-specific applications, the effectiveness of deformable modules and attention mechanisms has been validated in several enhanced YOLO variants. GFCD-YOLOXS incorporated the CBAM attention mechanism at the prediction layer, achieving 99.10% accuracy by guiding focus onto key lesion features despite environmental noise[31], while the GCS-YOLO model introduced a lightweight deformable convolution structure optimized for edge deployment without compromising precision[32]. The dual-branch design of Underwater-YOLO, although designed for marine use, serves as a compelling analog, wherein deformable convolutions and occlusion-aware attention mechanisms improved detection in dense, occluded settings, demonstrating potential cross-domain relevance[33]. Further, the fusion of local and global features, as shown in grape disease classification models using Swin Transformers and Group Shuffle Residual DeformNet, achieved accuracies above 98%, indicating that multi-scale deformable attention architectures are highly effective across varied plant pathology contexts[34].

The dual integration of deformable convolutions and advanced attention mechanisms, particularly as realized through the synergistic coupling of C2f-DCN and DPPA in the DCD-YOLOs model, reflects a significant evolution in plant disease detection frameworks. DPPA refinement of feature sensitivity—especially in mid-level representations—has been shown to enhance the recognition of subtle disease traits like the perforated edges of rotten-hole lesions and early speck patterns, which often blur with healthy leaf textures. This aligns with evidence from MSAM-YOLO, where the use of multi-scale attention allowed for a 4% improvement in identifying small, ambiguous grape leaf lesions by emphasizing multiscale textural granularity in diseased regions[35]. In related efforts, GFCD-YOLOXS embedded CBAM at the prediction end, which guided the model to focus on critical lesion features while suppressing environmental noise, achieving an impressive 99.10% accuracy[36]. This functional enhancement is further reinforced in EMA-YOLO, where efficient

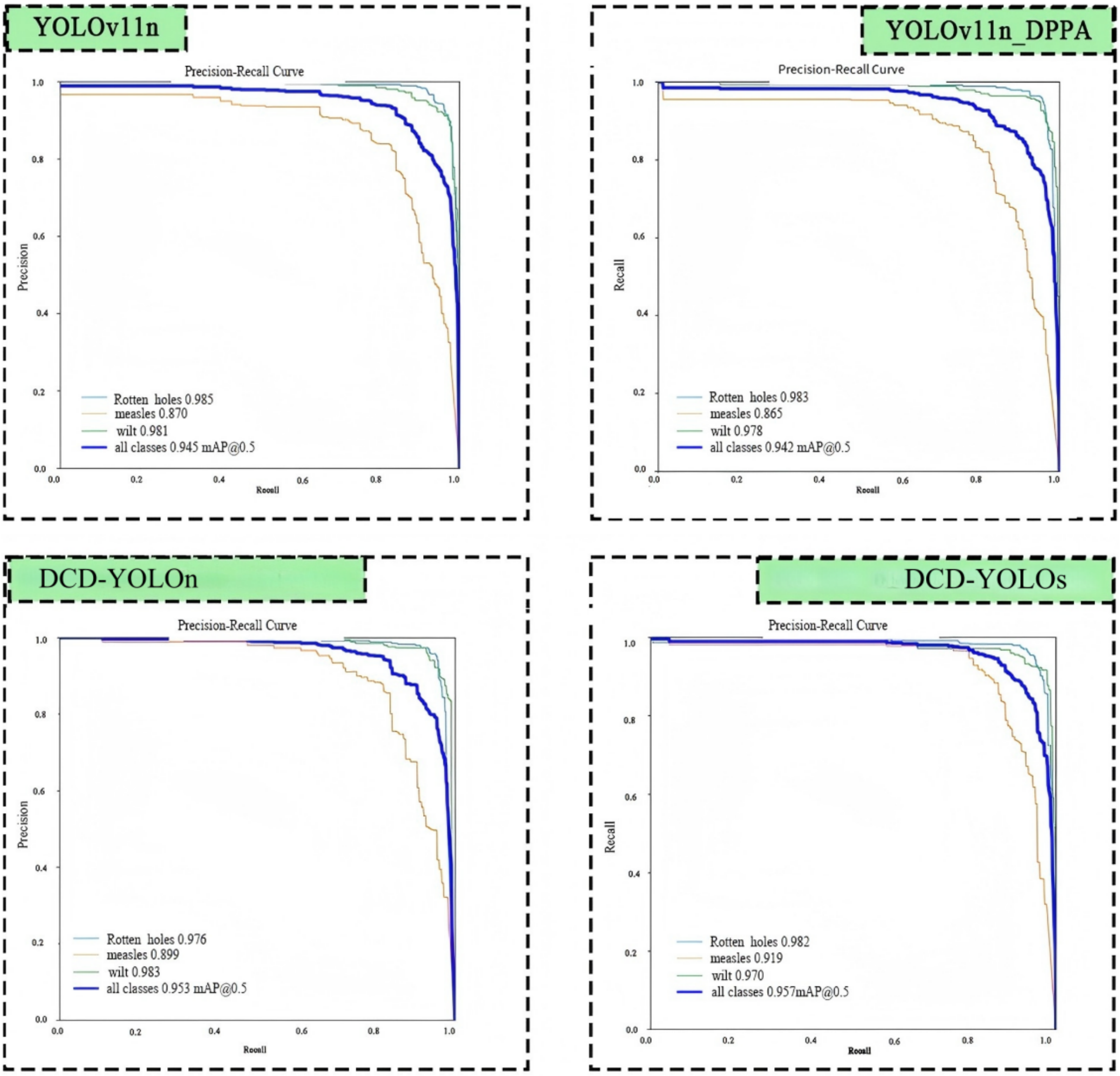

Fig. 9 Precision recall curves of the detection algorithm for three types of grape leaf diseases.

multi-scale attention modules improved the interpretability and recognition of black rot spots under resolution and lighting constraints, raising AP to 96.17%[37]. Similarly, GCS-YOLO combined Ghost modules with CBAM for computational efficiency without compromising fine-grained lesion detection, obtaining a mAP of 96.2% in grape leaf datasets[38]. A global-to-local attention strategy also showed promise in enhancing contextual prioritization, as evidenced by Underwater-YOLO occlusion-handling attention layers that improved object-background discrimination in densely packed environments, offering a comparable structural approach in viticulture settings[39]. Such dual-path attention designs have also been extended through Fusion Transformer YOLO, where real-time transformer modules fused with positional embeddings enhanced micro-lesion detection across complex scenes[40], and GSRDN-Swin Transformer networks, where contextual and deformable modules cooperatively yielded 98.6% classification accuracy across grape leaf disease datasets[41]. Notably, attention-enhanced YOLOv6 and CD-YOLO models applied CBAM and EMA variants to improve lesion saliency in cross-crop applications, confirming generalizability beyond viticulture[42]. These findings converge on a critical consensus: DPPA modules enrich feature semantics by prioritizing disease-relevant traits while suppressing background noise, and when harmonized with the flexible receptive fields of C2f-DCN, jointly facilitate robust lesion localization and classification, even under early symptom invisibility or occlusion, thus validating their coordinated utility in advanced smart agriculture frameworks.

The robustness of disease detection models in complex, uncontrolled field conditions remains a key determinant of their practical utility in precision agriculture. This study's verification of DCD-YOLOs under conditions of strong illumination and occlusion aligns with a growing body of research focused on enhancing YOLO-based architectures to address similar challenges. For instance, CEFW-YOLO employed multi-level linear attention and cross-channel attention modules, achieving a 5.2% improvement in mAP while maintaining high adaptability under diverse lighting and occlusion scenarios in apple orchards[43]. Likewise, the BED-YOLO combination of deformable convolutions and efficient attention modules significantly increased detection performance under natural environmental constraints, enhancing mAP to 91.3% in tomato datasets collected in the field[44]. Grape-specific detection systems,

https://doi.org/10.48130/frures-0026-0024

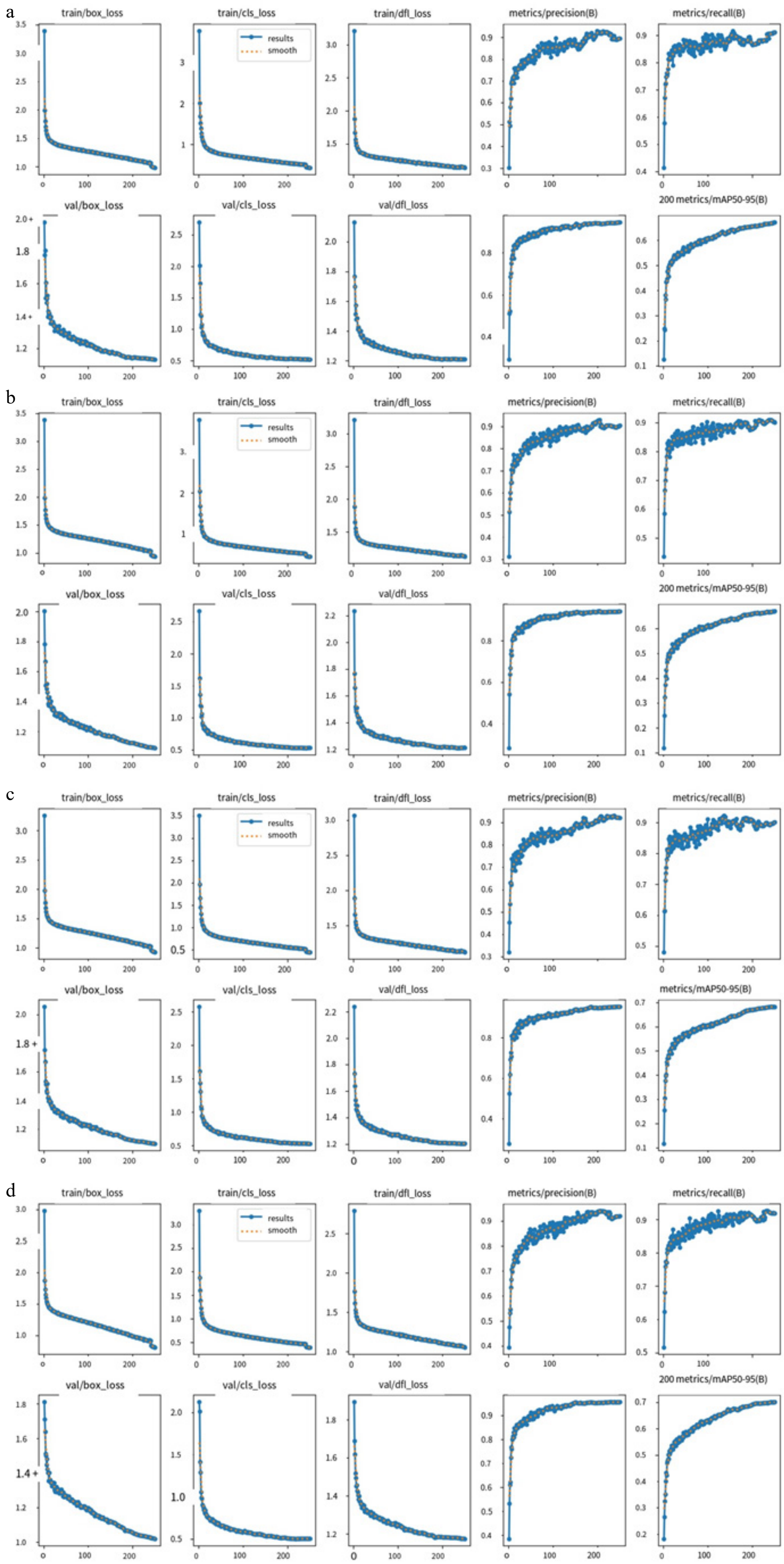

Fig. 10 Graphs generated by four algorithms for the detection of grape leaf diseases. (a) YOLOv11n. (b) YOLOv11n-DPPA. (c) DCD-YOLO. (d) DCD-YOLOs.

Table 4. Performance detection at different rounds.

Rounds	*P* (%)	*R* (%)	mAP (%)	Inference speed (ms)
100	0.838	0.896	0.922	5.0
150	0.861	0.899	0.929	4.7
200	0.911	0.878	0.943	5.0
250	0.923	0.901	0.953	4.8
300	0.908	0.906	0.941	4.9

Table 5. Detection results of different models.

Model	*P* (%)	*R* (%)	mAP (%)	Inference speed (ms)
DCD-YOLOs	0.920	0.919	0.957	9.8
DCD-YOLO	0.923	0.901	0.953	4.8
YOLOv11n-DPPA	0.905	0.902	0.942	4.8
YOLOv11n	0.896	0.909	0.945	3.5
YOLOv11n-DFF	0.862	0.921	0.941	11.6

including YOLO-GrapeNet and MSAM-YOLO, integrated multi-scale attention modules and advanced feature pyramids to address environmental noise and overlapping features, leading to mAP improvements exceeding 93% and 96%, respectively, under natural vineyard conditions[45]. Similar robustness was demonstrated in the DM-YOLO model for cucumber detection, where multi-context modules and global attention blocks enabled consistent performance under shadow and illumination changes[46]. The GFCD-YOLOXS model, using CBAM attention, and the GCS-YOLO architecture with lightweight modules, also maintained high accuracy (> 96%) while ensuring model efficiency in dynamic lighting conditions typical of vineyard settings[47]. Additional models, such as SlimFocal-YOLO and SPD-Conv-enhanced YOLOv8, demonstrated that refined backbone structures and attention integration can mitigate performance loss due to noise, further confirming the critical role of architectural robustness in deployment-ready agricultural detection systems[48]. Altogether, the consistent performance of these enhanced models under varying conditions reinforces the evidence that model robustness, enabled through refined attention mechanisms and contextual learning modules, is indispensable for real-world application in complex field environments.

The enhanced performance of our model in identifying measles diseases—characterized by fine, texture-blending lesions—can be directly attributed to the integration of DPPA and C2f-DCN modules, which collectively enable precise focus and contour adaptability. This improvement aligns with the results from MSAM-YOLO, a grape leaf detection model that employed a multi-scale attention module to enhance feature sensitivity for subtle, small targets, leading to a 4% accuracy improvement[49]. Similarly, BED-YOLO, applied to tomato leaves, combined deformable convolution with a multi-scale attention mechanism to improve recognition of blurred and occluded small lesions, increasing mAP to 91.3%, particularly excelling in challenging natural field images[50]. A related study utilizing CD-YOLO also emphasized attention and deformability for distinguishing fine disease patterns under complex backgrounds, yielding a 92.0% mAP[51]. In a different plant context, the YOLO-Tobacco model introduced hierarchical mixed-scale units and CBAM attention, substantially increasing detection precision of dense small lesion patterns in field-grown tobacco[52]. Likewise, CBAM-enhanced YOLOv6 models for general crop disease detection reported significantly greater diagnostic accuracy due to improved feature discrimination for localized, early-stage symptoms[53]. From a broader perspective, the YOLO-CXR model applied to chest X-rays demonstrated that attention-enhanced and lesion-focused architecture improved small lesion detection in cluttered visual contexts, an insight transferable to subtle leaf disease patterns[54]. In grape-specific detection, GFCD-YOLOXS introduced CBAM at the prediction stage, enabling the model to suppress noise and highlight minute disease traits in field-acquired images, achieving 99.10% accuracy[55]. Additionally, the YOLOv8-based model targeting grape leaf black rot integrated multi-scale attention with dilated convolutions to extract precise spatial features of minute lesions, attaining 96.17% AP despite image resolution constraints[56]. These corroborative results reinforce the critical role of adaptive attention and deformability in detecting fine-grained, visually ambiguous symptoms, validating our model design that strategically unifies global context focus (DPPA) and geometric adaptability (C2f-DCN) to address the complexities of early-stage disease detection.

In the context of deploying plant disease detection models on resource-limited edge devices, balancing accuracy with computational efficiency has emerged as a central challenge. The improved DCD-YOLO model demonstrated impressive mAP performance but still revealed trade-offs in inference speed compared to its lighter counterparts, reflecting a broader issue in current research. Numerous studies have investigated similar challenges and solutions through lightweight model architectures, pruning, and knowledge distillation. For example, SCD-YOLO, developed by incorporating slimming pruning and convolutional block attention mechanisms, significantly reduced parameter count while retaining high detection accuracy, proving suitable for real-time applications in edge computing environments[57]. Similarly, distillation-based strategies have shown potential in guiding pruned YOLO networks to retain performance, as seen in YOLOX-ViT for sonar object detection and M-YOLO-CRD for embedded systems, confirming the ability of student models to mimic larger networks under capacity constraints[58,59]. The need for model compression has prompted strategies such as pruning, quantization, and advanced knowledge distillation across different YOLO variants to meet the performance demands of real-time crop disease detection and industrial inspection scenarios[60,61]. Innovative methods such as the autoencoder-based KD and temperature-adaptive distillation further highlight the evolving sophistication of knowledge transfer mechanisms[62,63]. Such strategies become crucial when targeting scenarios like drone-based or robotic deployment in agricultural fields. Looking ahead, integrating multi-modal sensor data has been suggested as a promising avenue to enhance disease recognition beyond visual symptoms. Hyperspectral and thermal imaging offer complementary perspectives for detecting physiological stress and early-stage infections, though their fusion with deep learning frameworks remains a frontier of active research[64,65]. Beyond detection, the quantification of disease severity and its integration with predictive yield models holds transformative potential for precision agriculture, though few studies have robustly connected classification outcomes to actionable agronomic decisions. Therefore, future directions must emphasize lightweight model deployment, multimodal sensor fusion, and the transition from detection to decision support systems in real-world environments.

Traditional grape disease diagnosis methods rely on named, describable, spatially well-structured, and causally clear phenotypic features, with limitations such as poor timeliness, low accuracy, strong subjectivity, and difficulty in large-scale coverage. In contrast, AI-driven grape disease diagnosis methods depend on high-dimensional features of disease spots. The advantages include the ability to capture subtle differences imperceptible to the human eye, strong capability in extracting features of early-stage disease spots, consistent and repeatable diagnostic features, non-destructive sample handling, and rapid detection capabilities.

https://doi.org/10.48130/frures-0026-0024

In the future, with the construction of higher-quality open datasets, the deep integration of lightweight models and edge computing devices, and the refinement of the 'diagnosis-warning-decision' closed-loop system, AI-assisted diagnosis is expected to become a truly reliable 'digital grape protection expert' in the hands of grape growers.

Conclusions

The Grape Leaf Disease Detection dataset (GLDDD) constructed in this study contains 5,000 images of measles, wilt, rotten holes, and healthy leaves, which are divided into a training set, a validation set, and a test set in an 8:1:1 ratio. The category distribution is balanced and can effectively support the training and evaluation of the grape leaf disease detection model. The improvement strategy for the YOLOv11n model effectively enhanced the performance of grape leaf disease detection. By embedding the DPPA dual-channel attention mechanism module in the backbone network and fusing local details with global context features, the ability to discriminate heterogeneous spots such as diffuse spots of measles disease was enhanced. The C2f-DCN module was used instead of the original C3K2 structure, and deformable convolution was introduced to improve the localization accuracy of various morphologies such as perforation of rotten holes and irregular wilting.

The ablation test results showed that the C2f-DCN module performed well in improving model accuracy and mAP, and had little effect on parameter quantity and model size. The DPPA module can improve accuracy to some extent; the combination of the two can effectively enhance model performance, while the DFF module, although it can improve recall, will significantly increase the number of parameters and model size, and lead to a decrease in accuracy.

The improved DCD-YOLO model has the best performance, with an average accuracy of 95.7% for detecting grape leaf wilt, measles, and rotten holes, showing significant improvements compared to models such as YOLOv11n and YOLOv11n-DPPA. Especially for measles detection, the accuracy was 4.9% higher than that of YOLOv11n and 5.4% higher than that of YOLOv11n-DPPA. The model performs best at 250 iterations of training, when mAP reaches 95.3%, with precision and recall rates of 92.3% and 90.1%, respectively, achieving a good balance between precision and recall, which can meet the actual needs of grape leaf disease detection and provide effective technical support for the early diagnosis and control of grape diseases.

Ethical statements

This manuscript complies with the ethical standards and guidelines of the journal. All authors confirm that: (1) the research involves no human subjects, animal experiments, or sensitive data that require ethical approval; (2) there is no plagiarism, data fabrication, falsification, or other academic misconduct in the manuscript; (3) all materials, methods, and results are reported truthfully and accurately; (4) any potential conflicts of interest (financial, professional, or personal) have been fully disclosed in the manuscript; (5) appropriate permissions have been obtained for the use of any third-party data, figures, or content included in the work. We accept full responsibility for the ethical integrity of the research and the manuscript.

Author contributions

The authors confirm contributions to the paper as follows: study conception and design: Wang X, Wu Y; data collection: Wang X, Xiong Y; analysis and interpretation of results: Wang X, Yang L, Liao J, Wen Y; draft manuscript preparation: Wang X, Sheng S, Peng J, Li J, Wang S. All authors reviewed the results and approved the final version of the manuscript.

Data availability

These data were derived from the following resources available in the public domain: https://download.csdn.net/download/m0_64879847/88535108.

Acknowledgments

This research was supported by the National Technology Contract Management and Service Project (2025430102002140), the National Key Research and Development Program of China (2022YFD2200404), the Demonstration Project of 'Small Team' Science and Technology Aid to Xinjiang 'Expert Team for Arid Non-wood Forest Science and Technology Aid to Xinjiang from Central South University of Forestry and Technology' (2024-16-6), and the Central Financial Forestry Science and Technology Promotion Demonstration Fund Project (2023XT15).

Conflict of interest

The authors declare that tAll authors declare that there are no conflicts of interest that may affect the objectivity of this study in the design, implementation, data analysis, and paper writing process. Specifically, the author has no direct or indirect financial relationship with any commercial organization, non-profit organization, or individual (including but not limited to research funding, consulting fees, patent licensing fees, stock holdings, etc.); There is no personal or professional interest related to the research topic; The publication of research results does not involve any form of exchange of interests. The data in this study is authentic and reliable, and the analysis process adheres to the principle of scientific rigor. The conclusions are based solely on the research results and have not been influenced by any third-party interests. All authors have read and agree to the contents of this statement.hey have no conflict of interest.

Dates

Received 18 November 2025; Revised 7 April 2026; Accepted 22 April 2026; Published online 22 July 2026

References

[1] Rahman MU, Liu X, Wang X, Fan B. 2024. Grapevine gray mold disease: infection, defense and management. *Horticulture Research* 11:uhae182
[2] Zhu J, Wu A, Wang X, Zhang H. 2020. Identification of grape diseases using image analysis and BP neural networks. *Multimedia Tools and Applications* 79:14539−14551
[3] Wu A, Zhu J, He Y. 2019. Computer vision method applied for detecting diseases in grape leaf system. In *Cognitive Internet of Things: Frameworks, Tools and Applications*, ed. Lu H. vol 810. Cham: Springer. pp. 367−376 doi: 10.1007/978-3-030-04946-1_36
[4] Mousavi S, Farahani G. 2022. A novel enhanced VGG16 model to tackle grapevine leaves diseases with automatic method. *IEEE Access* 10:111564−111578
[5] University SPP, Wadekar NB, Sharma PK, Sapkale NS. 2015. Detection and controlling of grape leaf diseases using image processing and embedded system. *International Journal of Electrical and Electronics Engineering* 2:13−15

[6] Kolhalkar NR, Krishnan VL. 2020. Mechatronics system for diagnosis and treatment of major diseases in grape vineyards based on image processing. *Materials Today: Proceedings* 23:549–556
[7] Liu B, Ding Z, Tian L, He D, Li S, et al. 2020. Grape leaf disease identification using improved deep convolutional neural networks. *Frontiers in Plant Science* 11:1082
[8] Lu G, Zhang S, Que Y. 2026. Grapevine cultivation in the Xizang Plateau, the highest-altitude vineyard region: a systematic literature review. *Fruit Research* 6:e018
[9] Geng W, Gao X, Yang J, Zhang W, Xu F, et al. 2026. Chitosan and salicylic acid enhance postharvest resistance in grapes via integrated modulation of antioxidant, antifungal, and anti-senescence pathways. *Fruit Research* 6:e011
[10] Banerjee D, Sharma N, Upadhyay D, Dangi S, Yamsani N, et al. 2024. Hybrid CNN & random forest model for effective ginger leaf disease diagnosis. *2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM), Noida, India, 2024*. US: IEEE. pp. 1–6 doi: 10.1109/ICIPTM59628.2024.10563434
[11] Jadhav S, Yelapure S. 2024. Fuzzy rules in knowledge based diagnosis of anthracnose disease in grape crop. *International Journal for Multidisciplinary Research* 6:23750
[12] Carisse O, Morissette-Thomas V. 2013. Epidemiology of grape anthracnose: factors associated with defoliation of grape leaves infected by *Elsinoë ampelina*. *Plant Disease* 97:222–230
[13] Carisse O, Levasseur A, Provost C. 2021. Anthracnose risk establishment based on age-related susceptibility of grape leaves, flowers, and berries to infection by *Elsinoë ampelina*. *Plant Disease* 105:2380–2388
[14] Carisse O, Vincent S, Lafond-Lapalme J, Fall ML, Van der Heyden H. 2024. Quantitative insights into grapevine anthracnose (*Elsinoë ampelina*) epidemiology: impact of temperature and leaf age on incubation, lesion development, and sporulation. *Plant Disease* 108:2838–2844
[15] Yun HK, Louime C, Lu J. 2007. First report of anthracnose caused by *Elsinoe ampelina* on muscadine grapes (*Vitis rotundifolia*) in northern Florida. *Plant Disease* 91:905
[16] Chairma Lakshmi KR, Praveena B, Sahaana G, Nithya Jenev J, Gnanasekaran T, et al. 2023. YOLO for detecting plant diseases. *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2023*. US: IEEE. pp. 1029–1034 doi: 10.1109/ICAIS56108.2023.10073875
[17] Natij Y, El Karch H, Maafiri A, Mezouari A. 2024. Evaluating the performance of YOLO object detectors for plant disease detection. *2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM). Leeds, United Kingdom, 2024*. US: IEEE. pp. 1–6 doi: 10.1109/WINCOM62286.2024.10656918
[18] Ismail MA, Ali U, Ariyaluran Habeeb RA, Ali Shah SR. 2024. Performance evaluation of YOLO models in plant disease detection. *Journal of Informatics and Web Engineering* 3:199–211
[19] Ye R, Shao G, Yang Z, Sun Y, Gao Q, et al. 2024. Detection model of tea disease severity under low light intensity based on YOLOv8 and EnlightenGAN. *Plants* 13:1377
[20] Dai G, Fan J. 2022. An industrial-grade solution for crop disease image detection tasks. *Frontiers in Plant Science* 13:921057
[21] Ding Q. 2023. Crop disease recognition based on the YOLOv5 and transfer learning. *Applied and Computational Engineering* 13:46–57
[22] Kalezhi J, Shumba L. 2025. Cassava crop disease prediction and localization using object detection. *Crop Protection* 187:107001
[23] Sasmoko D, Siswanto E, Febryantahanuji F. 2025. Systematic literature review on CNN and YOLO algorithms for detecting plant diseases in precision agriculture. *Systematic Literature Review Journal* 1:12–20
[24] Iwano K, Shibuya S, Kagiwada S, Iyatomi H. 2024. Hierarchical object detection and recognition framework for practical plant disease diagnosis. *arXiv* 00:2407.17906
[25] Yu D, Wei W, Fan Z, Chen J, You Y, et al. 2023. VabHLH137 promotes proanthocyanidin and anthocyanin biosynthesis and enhances resistance to *Colletotrichum gloeosporioides* in grapevine. *Horticulture Research* 10:uhac261
[26] Wang X, Tu M, Wang Y, Zhang Y, Yin W, et al. 2024. Telomere-to-telomere and gap-free genome assembly of a susceptible grapevine species (Thompson Seedless) to facilitate grape functional genomics. *Horticulture Research* 11:uhad260
[27] Hrycan J, Forge T, Bowen P, Hart M, Úrbez-Torres JR. 2026. Investigating the interaction between ring nematode infestation and Petri disease development in young Merlot grapevines infected with *Phaeomoniella chlamydospora*. *European Journal of Plant Pathology* 174(5):823–840
[28] Ji M, Wu Z. 2022. Automatic detection and severity analysis of grape black measles disease based on deep learning and fuzzy logic. *Computers and Electronics in Agriculture* 193:106718
[29] Holkar SK, Bhondave SB, Patil SJ, Nanekar SC, Bhosale RK, et al. 2025. First report of Fusarium acutatum causing root rot and vascular wilt disease in grapevine (*Vitis vinifera*) in India. *Plant Disease* 109:2213
[30] Aubard M, Antal L, Madureira A, Ábrahám E. 2024. Knowledge distillation in YOLOX-ViT for side-scan sonar object detection. *arXiv* 00:2403.09313
[31] Yin X, Yuan L, Wang C, Ji T, Zhang Z, et al. 2026. Identification, characterization, and population structure of Colletotrichum spp. causing grape ripe rot in China. *Plant Disease* 110:944–956
[32] Chen Q, Xiong Q, Huang H, Tang S, Liu Z. 2024. Research on the construction of an efficient and lightweight online detection method for tiny surface defects through model compression and knowledge distillation. *Electronics* 13:253
[33] Ding JY, Zou CM, Jeon WS, Rhee SY. 2024. An improved YOLO detection approach for pinpointing cucumber diseases and pests. *Computers, Materials & Continua* 81:3989–4014
[34] Zhai C, Liu B, Li C, Zhao X, Liu H, et al. 2025. Research progress of detection and grading methods for major grapevine diseases. *Transactions of the Chinese Society for Agricultural Machinery* 56:341–359 (in Chinese)
[35] Dong Y, Li E, Jia Y, Zhuang Y. 2025. YOLO-GrapeNet: a robust model for grape disease detection. *2024 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Bangkok, Thailand, 2024*. US: IEEE. pp. 722–727 doi: 10.1109/WI-IAT62293.2024.00118
[36] Wang M, Li H, Wang K. 2025. The epidemiological patterns and management strategies of grape anthracnose. *China Fruit & Vegetable* 45:38–42 (in Chinese)
[37] Du XQ, Yang YL, Lu J. 2025. Screening and characterization of biocontrol *Bacillus* strains against *Colletotrichum* causing grape ripe rot. *Horticulturae* 11:802
[38] Haijoub A, Hatim A, Arioua M, Eloualkadi A, Gómez-López MD. 2024. Enhancing plant disease detection in agriculture through YOLOv6 integration with convolutional block attention module. *Bioinspired Systems for Translational Applications: From Robotics to Social Engineering*, eds Ferrández Vicente JM, Val Calvo M, Adeli H. Vol 14675. Cham: Springer. pp. 474–484 doi: 10.1007/978-3-031-61137-7_44
[39] Hao S, Li X, Peng W, Fan Z, Ji Z, et al. 2024. YOLO-CXR: a novel detection network for locating multiple small lesions in chest X-ray images. *IEEE Access* 12:156003–156019
[40] Hu Q, Zhang Y. 2025. GCS-YOLO: a lightweight detection algorithm for grape leaf diseases based on improved YOLOv8. *Applied Sciences* 15:3910
[41] Huang B, Ding Y, Liu G, Tian G, Wang S. 2024. ASD-YOLO: an aircraft surface defects detection method using deformable convolution and attention mechanism. *Measurement* 238:115300
[42] Holkar SK, Bhanbhane VC, Ghotgalkar PS, Markad HN, Lodha TD, et al. 2024. Characterization and bioefficacy of grapevine bacterial endophytes against *Colletotrichum gloeosporioides* causing anthracnose disease. *Frontiers in Microbiology* 15:1502788
[43] Karthik R, Vardhan GV, Khaitan S, Harisankar RNR, Menaka R, et al. 2024. A dual-track feature fusion model utilizing Group Shuffle Residual DeformNet and swin transformer for the classification of grape leaf diseases. *Scientific Reports* 14:14510
[44] Steiner DRM, Modesto LR, Dias AH, Zappelini J, Petters-Vandresen DAL, et al. 2025. *Colletotrichum nymphaeae* and *Colletotrichum theobromicola* isolated from anthracnose symptoms cause grape ripe rot. *Plant Pathology* 74:813–824

https://doi.org/10.48130/frures-0026-0024

[45] Li H, Zhuang X, Bao S, Chen J, Yang C. 2024. SCD-YOLO: a lightweight vehicle target detection method based on improved YOLOv5n. *Journal of Electronic Imaging* 33:023041

[46] Li T, Zhang L, Lin J. 2024. Precision agriculture with YOLO-Leaf: advanced methods for detecting apple leaf diseases. *Frontiers in Plant Science* 15:1452502

[47] Li Z, Zheng B, Chao D, Zhu W, Li H, et al. 2024. Underwater-YOLO: underwater object detection network with dilated deformable convolutions and dual-branch occlusion attention mechanism. *Journal of Marine Science and Engineering* 12:2291

[48] Lin J, Chen T, Liu X, Chen Z, Lei Y. 2025. Salicylic acid represses VdMYB31 expression to enhance grape resistance to *Colletotrichum viniferum*. *International Journal of Biological Macromolecules* 288:138731

[49] Liu B, Jiang W. 2024. DFKD: dynamic focused knowledge distillation approach for insulator defect detection. *IEEE Transactions on Instrumentation and Measurement* 73:5039916

[50] Liu R, Zeng W. 2026. Automatic detection of structural defects in tunnel lining via network pruning and knowledge distillation in YOLO. *Structural Health Monitoring* 25(2):1165−1181

[51] Liu Y, Yu Q, Geng S. 2024. Real-time and lightweight detection of grape diseases based on Fusion Transformer YOLO. *Frontiers in Plant Science* 15:1269423

[52] Bai Q, Chen D, He S, Wang Y, Chen X, et al. 2024. Identification of the pathogen causing grape white rot in Fangshan District, Beijing, and preparation of a rapid sporulation medium. *Journal of Plant Protection* 51:1327−1334 (in Chinese)

[53] Ren J, Zhang H, Wang G, Dai C, Teng F, et al. 2024. Real-time grape disease detection model based on improved YOLOv8s. *INMATEH Agricultural Engineering* 72:96−105

[54] Seo Y, Kang JK, Kim Y. 2024. Autoencoder-based knowledge distillation for quantized YOLO detector. *2024 21st International SoC Design Conference (ISOCC), Sapporo, Japan, 2024*. US: IEEE. pp. 376−377 doi: 10.1109/ISOCC62682.2024.10762165

[55] He Y, Liu X, Wu L, Jiang P, Ma H, et al. 2023. Molecular characterization of a glycosyltransferase gene TaUGT7 and its resistance to Fusarium head blight. *Journal of Plant Protection* 50:197−205 (in Chinese)

[56] Setyanto A, Bayu Sasongko T, Fikri MA, Ariatmanto D, Agastya IMA, et al. 2025. Knowledge distillation in object detection for resource-constrained edge computing. *IEEE Access* 13:18200−18214

[57] Tian Y, Yao J, Sheng C, Lv M, Kang Q. 2024. SlimFocal-YOLO: a lightweight grape disease detection model based on YOLOv8. *DSAI '24: Proceedings of the 2024 International Conference on Digital Society and Artificial Intelligence*. Qingdao China: ACM. pp. 434−440 doi: 10.1145/3677892.3677959

[58] Wang C, Wang Y, Ma G, Bian G, Ma C. 2023. Identification of grape diseases based on improved YOLOXS. *Applied Sciences* 13:5978

[59] Wang Q, Yan N, Qin Y, Zhang X, Li X. 2025. BED-YOLO: an enhanced YOLOv10n-based tomato leaf disease detection algorithm. *Sensors* 25:2882

[60] Ye Z, Wang Z, Chen H, Wang W. 2025. A lightweight rice disease object detection algorithm improved on YOLOv8. *2024 12th International Conference on Information Systems and Computing Technology (ISCTech), Xi'an, China, 2024*. US: IEEE. pp. 1−8 doi: 10.1109/ISCTech63666.2024.10845418

[61] Yi Y, Song Y, Qin J, Gao F, Han H. 2025. CD-YOLO: an enhanced YOLOv8n-based algorithm for accurate crop disease detection. *2025 IEEE 8th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 2025*. US: IEEE. pp. 285−290 doi: 10.1109/ITOEC63606.2025.10968360

[62] Zhu J, Qiu J, Chen S, Chen S, Zhang H. 2025. An application of YOLOv8 integrated with attention mechanisms for detection of grape leaf black rot spots. *PLoS One* 20:e0321788

[63] Chen S, Jiao R, Li Z, Liu Z, Li G. 2022. Identification and biological characteristics of Cladosporium pathogens causing grape fruits rot in Xinjiang. *Plant Protection* 48:180−187,202

[64] Zhang G, Liu S, Nie S, Yun L. 2024. YOLO-RDP: lightweight steel defect detection through improved YOLOv7-tiny and model pruning. *Symmetry* 16:458

[65] Donegan MA, Kahn AK, Becker N, Castillo Siri A, Campos PE, et al. 2025. Century-old herbarium specimen provides insights into Pierce's disease of grapevines emergence in the Americas. *Current Biology* 35:145−153.e4