Smart Agricultural Technology 12 (2025) 101433

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

RSD-YOLO: An improved YOLOv7-tiny framework for oat disease severity identification with integration of ReXNet and decoupled head

Yongquan Zhang [a], Yiwei Xu [a,f], Taosheng Xu [b], Changmiao Wang [c], Chengdao Li [d], Hai Wang [e,f,*]

[a] School of Data Science, Zhejiang University of Finance and Economics, Hangzhou, 310018, China
[b] Institute of Intelligent Machines, Hefei Institutes of Physical Science, Chinese Academy of Science, Hefei, 230031, China
[c] Shenzhen Research Institute of Big Data, Shenzhen, 518172, China
[d] Food Futures Institute, Murdoch University, Perth, WA6150, Australia
[e] School of Engineering and Energy, Murdoch University, Perth, WA6150, Australia
[f] Harry Butler Institute, Murdoch University, Perth, WA6150, Australia

ARTICLE INFO

Keywords:
YOLOv7-tiny
Oat crop diseases
Oat disease severity identification
Regularized Xception-based network
Decoupled head structure

ABSTRACT

In precision agriculture, accurate and timely identification of plant disease severity is essential for optimizing crop yield and health. However, current methods often face challenges such as high computational cost and reduced accuracy in resource-constrained environments, limiting their practical use on farms. To address these limitations, we propose RSD-YOLO, an improved YOLOv7-tiny model that integrates a Regularized Xception-based Network (ReXNet), a Slim-Neck module, and a Decoupled Head—together forming the RSD design. We construct a dataset of 1,010 oat leaf images, categorized into five severity levels and annotated by experts. RSD-YOLO achieves 91.6% precision, 90.8% recall, and 88.5% mAP@0.5, significantly outperforming YOLOv7-tiny by up to 10%, while maintaining a computational cost of only 11.2 GFLOPs. Recent studies have applied lightweight models such as EfficientSAM and SwiftFormer for crop health monitoring on drones and edge devices. However, these models often struggle to balance accuracy and efficiency. In contrast, RSD-YOLO achieves higher performance with lower computational cost, making it well-suited for real-time deployment in agricultural environments.

1. Introduction

Oats are among the world's most vital crops, contributing significantly to nutrition and food production [1,2]. They are also widely used in animal feed and crop rotation [3]. In 2022, global oat production reached approximately 26 million metric tons, underlining their importance in the agricultural industry [4]. However, oat cultivation faces significant challenges, particularly from diseases like leaf rust and neck rust, which can severely reduce both yield and quality, with potential losses as high as 85% [5]. Thus, accurately identifying the severity of oat diseases, is a critical challenge in precision agriculture. Current disease detection methods, such as traditional visual assessments, are highly subjective and often unreliable due to disease similarities and environmental influences [6]. These methods typically require experienced specialists, making timely, accurate identification difficult, especially in large-scale agricultural operations. Hyperspectral imaging has also been used for disease severity assessment, but this requires expensive equipment and technical expertise, limiting its feasibility for widespread use [7]. These challenges highlight the urgent need for an automated, cost-effective, and scalable solution for identifying disease severity in oat crops.

Recent advancements in computer vision have enabled more effective automated plant disease detection. Techniques such as deep learning have been successfully applied to tasks like plant species classification [8–10], leaf disease detection [11–13]. Unlike traditional methods, deep learning models can automatically extract features from images, eliminating the need for complex preprocessing and enabling end-to-end detection. Recent advances in crop disease monitoring have leveraged both terrestrial and aerial imaging systems for disease identification and classification [14]. Terrestrial imaging—typically involving handheld cameras or ground-mounted sensors—provides high-resolution images that support leaf- or plant-level disease identification. For instance, Karim et al. (2024) developed a lightweight MobileNetV3 model deployed on an edge device to classify grape leaf diseases in real-

* Corresponding author at: School of Engineering and Energy; Harry Butler Institute, Murdoch University, Perth, WA6150, Australia.
E-mail address: Hai.Wang@murdoch.edu.au (H. Wang).

https://doi.org/10.1016/j.atech.2025.101433
Received 9 May 2025; Received in revised form 8 September 2025; Accepted 9 September 2025

Available online 16 September 2025

time, achieving over 99% accuracy using Grad-CAM explanations [15]. Shruthi et al. (2024) combined MobileNetV3 with wavelet-kernel extreme learning machine for multi-disease classification across eight crops, demonstrating robust performance on over 20K leaf images [16]. These methods benefit from rich spatial detail but often struggle with variable lighting, background clutter, and leaf occlusion in field settings. By contrast, aerial imaging, especially via UAVs, enables rapid monitoring over broad areas. Shahi et al.'s (2023) survey highlights the effectiveness of UAV platforms for disease detection using RGB, multispectral, and hyperspectral sensors [17]. Lyu et al. (2025) demonstrated the use of lightweight ViT-based models with UAV-acquired RGB imagery to detect downy mildew in kimchi cabbage, achieving up to 94.8% accuracy [18]. Though aerial data supports efficient large-area surveillance, its typical resolution and canopy occlusion hinder fine-grained lesion analysis.

However, both image acquisition methods face critical limitations in tasks requiring fine-grained severity assessment. Aerial imagery suffers from lower spatial resolution and canopy occlusion, making it difficult to capture subtle lesion patterns. Terrestrial field photography, though higher in resolution, is often affected by uncontrolled lighting, background clutter, leaf overlapping, and non-uniform leaf orientation—all of which reduce the consistency and precision of visual features required for reliable disease severity quantification. These challenges are especially pronounced in oat crops, whose narrow, elongated leaves with dense parallel venation present unique difficulties for image-based severity analysis. Disease symptoms on oat leaves, such as those caused by leaf rust or crown rust, are often subtle and aligned along the venation, making them difficult to detect without high-contrast, clean imaging conditions. Additionally, the erect growth habit and overlapping foliage of oat plants frequently lead to occlusion and shadowing in field images.

To overcome these limitations, this study adopts a controlled indoor imaging protocol. Oat leaves are photographed individually against a pure white background using consistent lighting, ensuring that lesion boundaries and severity features are clearly visible and minimally affected by environmental noise. This setup allows for more precise expert annotation and provides higher-quality input data for model training. By focusing on laboratory-acquired imagery, the proposed approach addresses the shortcomings of traditional field imaging systems and enables a more accurate and scalable solution for oat disease severity identification.

Despite the promising results of deep learning in disease detection, its application to disease severity estimation remains limited. Early work by Wang et al. (2017) Convolutional Neural Networks (CNNs) were first applied to classify apple black rot images [19]. However, this approach was specific to certain diseases and lacked broader applicability. Similarly, Su et al. (2020) combined ResNet-101 with semantic segmentation for predicting wheat scab, but the accuracy of 77.19% indicated room for improvement [20]. PD2SE-Net proposed by Liang et al. (2019) introduced PD2SE-Net, a framework that performed both disease classification and severity estimation [21]. Although it achieved high precision, the model's complexity resulted in increased computational demands, restricting its deployment on edge devices or drones commonly used in modern precision agriculture. In later studies, BLSNet and SegNet demonstrated notable advancements in plant disease severity estimation through semantic segmentation techniques [22,23]. BLSNet was specifically designed for the recognition and segmentation of bacterial leaf streak lesions in rice leaves. Built upon a UNet architecture, it incorporated an attention mechanism and multi-scale feature extraction to enhance lesion segmentation accuracy. SegNet focused on detecting late blight lesions in potatoes caused by the pathogen Phytophthora infestans. The model processed high-resolution Red-Green-Blue images, and achieved high intersection over union (IoU) scores for background and disease lesions, respectively. Esgario et al. (2020) proposed a multi-task system based on ResNet50 for identifying and estimating the severity of biotic stress on coffee leaves [24]. Their model used data augmentation to enhance robustness and accuracy. The computational experiments achieved 86.27% for severity estimation. Fenu and Malloci (2021) introduced a multi-output deep learning system to classify three diseases and six severity levels in pear leaves [25]. The model was evaluated on the DiaMOS Plant dataset, achieving 78.31% for severity estimation. However, Both models depend on extensive labeled datasets like the DiaMOS Plant dataset, which can be costly and labor-intensive to gather, particularly for emerging crops and diseases. Prabhakar et al. (2020) studied early blight in tomato leaves using ResNet101 trained on the PlantVillage dataset [26]. Zeng et al. (2020) addressed the challenge of detecting severity levels in Huanglongbing (HLB)-infected citrus leaves. Nevertheless, both studies required large datasets and high-resolution images, making their approaches computationally expensive and difficult to deploy in real-time applications or on low-resource devices [27]. Goncalves et al. (2021) focused on severity estimation using semantic segmentation, achieving reliable results across multiple crop diseases. However, their method relied heavily on large-scale datasets and high-resolution images, posing challenges for scalability and adaptability in practical farming environments [28]. These challenges emphasize needing effective models to accurately evaluate disease severity in crops like oats, where subtle symptoms are hard to detect. The aim is to develop a model that achieves high accuracy with low computational demands, suitable for real-time use in small to medium-sized farms.

To address the aforementioned challenges, this paper builds upon YOLOv7-tiny [29] and integrates the Decoupled Head structure from YOLOv11 [30], along with a Regularized Xception-based Network (ReXNet) [31] and Slim-Neck [32], to propose an improved model, RSD-YOLO. The model uses a ReXNet to improve feature extraction and lower computational complexity by optimizing channel dimensions between layers. The Slim-Neck module is incorporated to minimize computational costs further. By using the Decoupled Head structure from YOLOv11, the model clearly distinguishes between classification and bounding box regression, enhancing both identification and positioning accuracy. This combined design specifically optimizes the model for identifying the severity of oat diseases in resource-constrained environments. The main contributions of this paper are as follows:

(i) A novel backbone-fusion architecture is proposed by integrating ReXNet with a Slim-Neck module. ReXNet optimizes channel-wise feature representation, enabling the model to capture subtle lesion patterns under complex textures, while Slim-Neck ensures efficient multi-scale feature fusion. This design jointly enhances detection accuracy and computational efficiency.
(ii) The model adopts the Decoupled Head from YOLOv11, which separates classification and localization tasks into parallel branches. This task-specific optimization reduces conflict between objectives, thereby improving detection robustness in challenging conditions such as uneven lesion distribution or subtle disease symptoms.
(iii) A comprehensive and well-annotated oat leaf disease severity dataset is constructed, categorized into five severity levels. Expert-validated labels and rich data augmentation ensure reliable model training. This dataset provides a foundation for further research in crop disease severity estimation.
(iv) Experimental results demonstrate that RSD-YOLO improves precision by 7.0%, recall by 10.0%, and mAP@0.5 by 7.6% compared to the YOLOv7-tiny baseline, while maintaining a significantly lower computational cost than heavier models like Faster R-CNN or DEtection TRansformer (DETR). Studies validate the efficiency of its main features, ReXNet, Slim-Neck, and Decoupled Head, demonstrating its suitability for resource-constrained settings.

The remainder of this paper is structured as follows: Section 2 discusses dataset creation and augmentation, introduces the RSD-YOLO model, and describes how ReXNet, Slim-Neck, and Decoupled Head modules enhance identification efficiency and accuracy. Section 3 discusses the experimental results, including performance comparisons

Fig. 1. Oat plants with disease in the glasshouse at DPIRD, Perth, Australia.

Fig. 2. Dataset example. H: Healthy, R: Resistant, MR: Moderately Resistant, MS: Moderately Susceptible, S: Susceptible.

with other YOLO models and an ablation study on key components. We also present that, as evidenced by experimental results and architectural analysis, the model has key advantages and limitations. Section 4 concludes the article with a summary of findings and potential directions for future work.

2. Materials and methods

2.1. Dataset

2.1.1. Dataset acquisition

The images used in this study were captured from oat plants grown in pots, as shown in Fig. 1. The oat variety, PALESTINE, was provided by the Western Crop Genetics Alliance in Australia. The inoculation of the plants was conducted in glasshouse 40.4 of the Cereal Pathology Research Group, located at the Department of Primary Industries and Regional Development in South Perth, Australia.

For data collection, the images were taken using an iPhone 12 Pro Max, ensuring high-resolution image quality. The controlled glasshouse environment provided consistent lighting conditions, which helped minimize variations in image quality and ensured clear visibility of disease symptoms on oat leaves. This approach ensured the reliability and uniformity of the dataset, which is critical for accurate disease severity identification and analysis.

2.1.2. Dataset construction

As shown in Fig. 2, the dataset was developed to evaluate the severity of disease in oat leaves and is categorized into five classes based on visual symptoms: Healthy, Resistant, Moderately Resistant, Moderately Susceptible, and Susceptible. This five-class scheme is informed by standard visual diagnostic protocols commonly used in plant pathology. The categories reflect a gradual progression of structural degeneration in leaf appearance, ranging from uniform greenness to increasing levels of discoloration, lesion spread, and tissue deformation. This classification enables the model to learn subtle visual cues related to lesion severity, distribution, and the extent of damage. Specifically:

- **Healthy (H):** Leaves exhibit no signs of disease, displaying a uniform green color without any visible lesions or discoloration.
- **Resistant (R):** Leaves show minor disease symptoms but retain overall structural integrity and green coloration. Symptoms are minimal and localized.
- **Moderately Resistant (MR):** Leaves exhibit moderate disease symptoms, including some discoloration and slight deformation. Lesions are more noticeable but do not greatly affect the leaf's structure.
- **Moderately Susceptible (MS):** Leaves display clear symptoms like lesions, discoloration, and structural damage. Brown or yellowish areas partially replace the green color.
- **Susceptible (S):** Leaves are severely affected by disease, displaying extensive damage, discoloration, and deformation. The structure is seriously weakened, showing widespread browning and shriveling.

The development of the disease is usually a gradual process, from healthy to slightly damaged, to moderate and severe infection, and finally to irreversible severe damage. This classification method illustrates the disease's progression, aiding farmers and researchers in understanding its development and forming effective management strategies. The dataset comprises a total of 1,010 oat leaf images divided into five severity categories: healthy (H, 130), resistant (R, 468), moderately resistant (MR, 432), moderately susceptible (MS, 343), and susceptible (S, 70). All images were captured under controlled lighting conditions to ensure consistency and minimize variations during data acquisition.

Although most existing studies use whole plants as datasets, they usually focus on studying the spatial distribution of diseases, disease transmission patterns, and overall growth status assessment. The research object of this paper is the identification of disease severity in oat leaves. Leaf datasets are easier to accurately annotate, such as the location, size, and morphology of lesions, which is conducive to training high-precision models for detection and classification. Therefore, leaf image datasets are selected instead of whole plant images.

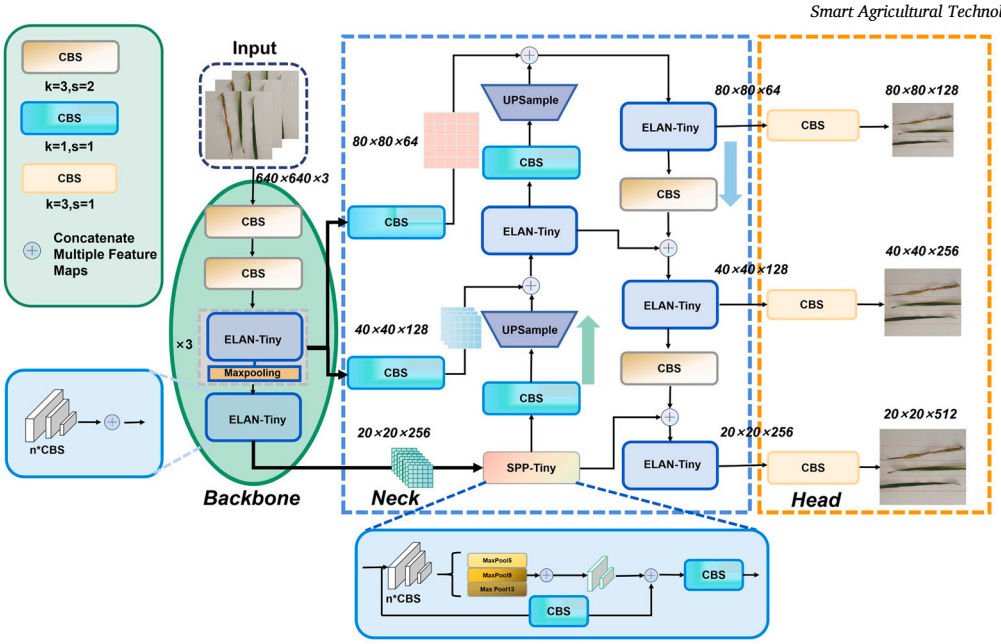

Fig. 3. Model optimization framework. Backbone: Extract high-level features. Neck: Generate multi-scale features. Head: Output detection result.

2.1.3. Data processing

To ensure the dataset accurately represents the varying degrees of oat leaf disease severity, strict inclusion criteria are applied. Images are collected from experimental fields and greenhouses, focusing on five disease severity categories: H, R, MR, MS, and S. The classification is based on visible symptoms such as discoloration, lesion size, and structural damage. Only high-resolution images with clear disease symptoms and minimal background noise are included. Images with blurriness, poor lighting, or occlusions are excluded. Leaves are placed on a white background during image capture to reduce noise.

To address class imbalance, we adopt class-targeted oversampling with augmentation for the minority categories and then apply a unified augmentation pipeline to all classes. Concretely, each image in the H class is augmented twice and each image in the S class is augmented four times respectively. The class-targeted augmentations included rotation, flip, mild scaling, translation, light color jitter, and Gaussian blur, ensuring shape- and lesion-preserving variations.

The dataset is annotated using LabelImg, following a structured workflow. Each image is manually labeled with its respective category, and bounding boxes are drawn around disease-affected regions when applicable [33]. Annotations are saved in YOLO-compatible format [34]. To ensure accuracy, all labeling is reviewed by plant pathology experts. The dataset is then split into training (80%), validation (10%), and test (10%) subsets, ensuring a balanced distribution of disease severity categories.

After balancing and annotating, to enhance model generalization and simulate diverse field conditions, data augmentation techniques are applied. Augmentation operations involve random rotations of $\pm 90^{\circ}$ or 180°, scaling factors of 0.50, 0.75, 1.00, and 1.25, brightness adjustments from 0.50 to 1.50, and random horizontal and vertical flips. These transformations are implemented using libraries such as Albumentations, and annotations are adjusted accordingly to maintain alignment with the augmented images [35].

Quality control is performed after augmentation to ensure the integrity of the dataset. All augmented images and corresponding labels are manually inspected to eliminate errors or inconsistencies. This preparation process makes the dataset strong, varied, and suitable for training models to identify oat leaf disease severity.

2.2. Overview of YOLOv7-tiny

As shown in Fig. 3, YOLOv7-tiny is a compact and efficient object detection model within the YOLO family [29]. It is specifically designed for scenarios with limited computational resources. This model extends the YOLOv7 architecture, which is renowned for its balance between accuracy and speed in real-time object detection. The YOLOv7-tiny model architecture consists of Backbone, Neck, and Head.

Backbone: Employs CSPDarkNet53 for feature extraction, boosting efficiency through Cross-Stage Partial connections. Convolutional layers and Leaky ReLU activation functions process the input image, and downsampling to obtain high-level feature representations.

Neck: Links the backbone and head, using a Path Aggregation Network and Feature Pyramid Network to enhance the detection of objects at different scales.

Head: Convert features into detection results, providing bounding box coordinates, object classes, and confidence scores. It uses predefined anchor boxes for bounding box prediction and performs regression and classification for object detection.

Together, these components make YOLOv7-tiny efficient for devices with limited computational capabilities.

2.3. Novel network construction

To enhance the detection accuracy and efficiency of the YOLOv7-tiny framework under resource constraints, we propose a novel lightweight model architecture RSD-YOLO specifically designed for oat disease severity identification. As shown in Fig. 4, our model introduces three key innovations: a Regularized Xception-based Network (ReXNet) backbone, a Slim-Neck structure built with GSConv and VoV-GSCSP, and a Decoupled Head inspired by YOLOv11. The integration of these components significantly improves multi-scale feature learning and classification-localization decoupling, while keeping the computational complexity low.

2.3.1. Regularized Xception-based network (ReXNet)

In our framework, we innovatively incorporate Regularized Xception-based Network (ReXNet). ReXNet is designed to tackle the challenges of developing efficient, high-performance neural networks, particularly for image classification tasks. Its architecture focuses on optimizing channel

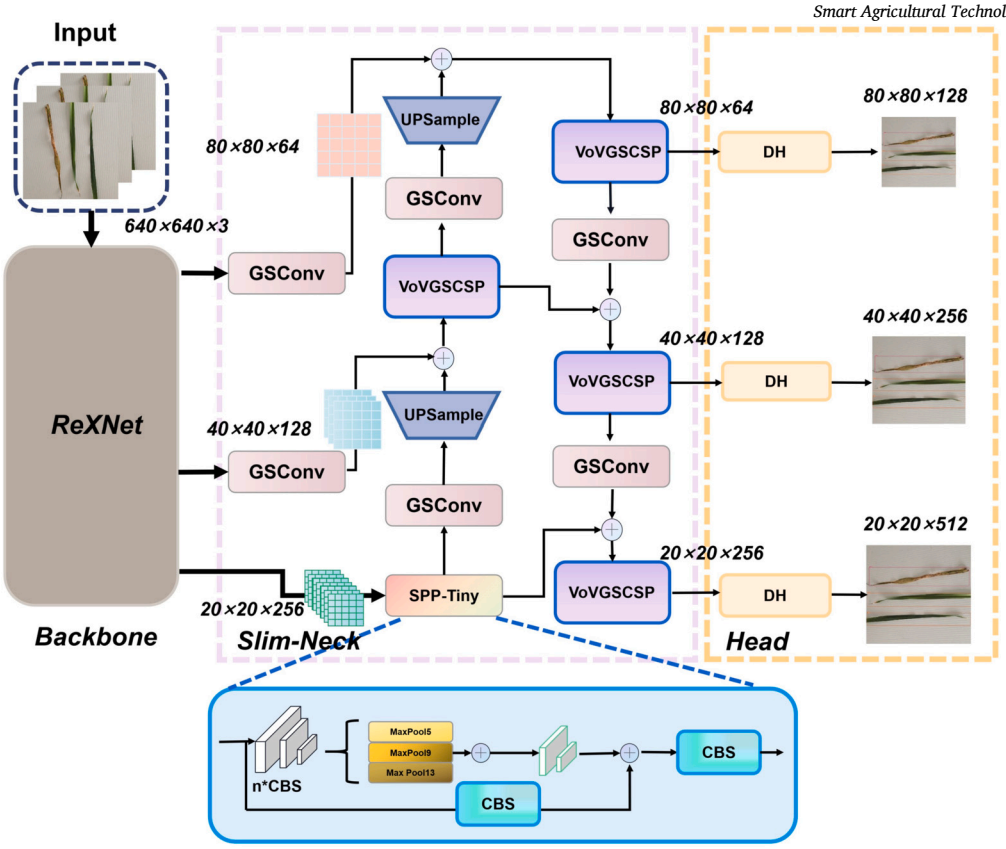

Fig. 4. Construction of our model. ReXNet Backbone: Optimize channel configuration and extract features. Slim-Neck: Simplify feature fusion structure and improve processing efficiency, The structure of GSConv and VoV-GSCSP can be seen in Fig. 5 below. Decoupled Head (DH): Separate target classification and positioning and output the final detection results. The structure of DH can be seen in Fig. 6 below.

dimensions across the network's layers, moving away from conventional methods that may limit model expressiveness due to computational constraints.

Unlike traditional models like MobileNetV2, which use a stage-wise channel configuration that gradually expands channels, ReXNet introduces a novel approach by parameterizing channel dimensions as a linear function of the block index. This linear increase in output channel dimensions across the network has been empirically shown to enhance accuracy. The ReXNet channel dimension formula is as follows:

$$c_i = a f(i) + b, \tag{1}$$

where c_i represents the output channel of the $i-th$ block, $f(i)$ is a piecewise linear function determined by the block index, and a and b are parameters adjusted during the network search process to optimize performance.

The core of ReXNet's structure involves inverted bottleneck blocks with depthwise separable convolutions (DSC). This design reduces the number of parameters and computational cost while maintaining the ability to capture complex features. The deep convolution formula in ReXNet is:

$$Y = \sigma(W_d \cdot (ReLU(W_p \cdot X) + b_p)) + b_d, \tag{2}$$

where X is the input, W_p and b_p are the weights and biases for the pointwise convolution, W_d and b_d are for the depthwise convolution, and σ is the activation function, often $ReLU$ or its variants.

It also integrates Squeeze-and-Excitation (SE) modules to dynamically recalibrate channel importance. These blocks enhance the network's ability to generalize across different tasks. The dynamic adjustment process of the channels in the SE module is:

$$s = \sigma(W_2 \cdot ReLU(W_1 \cdot GlobalAvgPool(E))), \tag{3}$$

where E is the feature map, and s scales the output channels.

Diseases on oat leaves usually appear as low-contrast brown or yellow spots, which are difficult to recognize by traditional convolutional networks. ReXNet enhances the ability to extract subtle features by adjusting the channel configuration and using an inverted bottleneck structure, enabling the model to capture these subtle disease features. The surface of oat leaves has a complex texture structure, and disease symptoms are often distributed along the veins. ReXNet uses DSC, which not only reduces the computational cost but also better extracts the local texture features of the leaves, thereby improving the identification accuracy of the diseased severity.

2.3.2. Slim-neck

As shown in Fig. 5 (a), Ghost Shuffle Convolution (GSConv) is a lightweight convolution operation designed to improve the efficiency of deep neural networks, especially in real-time detection tasks. GSConv combines the advantages of standard convolution and DSC, reducing computational complexity through hybrid features while retaining rich feature expression capabilities.

Based on GSConv, VoVNet Ghost Shuffle Cross Stage Partial (VoV-GSCSP) further optimizes the efficiency through a cross-stage partial network (CSP) architecture, as shown in Fig. 5 (b). This design maintains high accuracy while reducing computational cost and inference time. GSConv incorporates a bottleneck module designed to decrease the input dimension and subsequently restore it, optimizing the model's efficiency while preserving essential features.

The Slim-Neck module enhances the model's ability to fuse features across multiple scales, allowing for better adaptation to disease areas of varying sizes and shapes, which in turn improves the recognition of complex disease patterns. As illustrated in Fig. 5, the standard convolution and ELAN-tiny modules are replaced with GSConv and VoV-GSCSP,

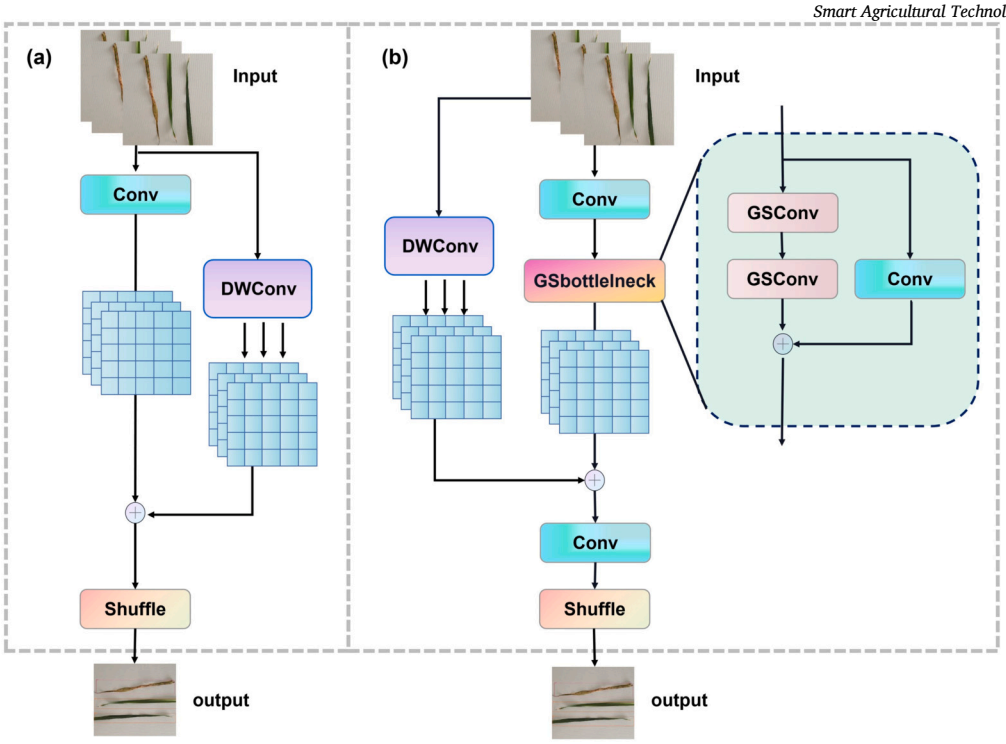

Fig. 5. The structure of GSConv (a) and VoV-GSCSP (b). (a) GSConv utilizes a combination of standard convolution and DWConv to enhance feature extraction efficiency. (b) VoV-GSCSP integrates GSConv into a CSP architecture, optimizing computational efficiency while maintaining robust feature representation. Shuffle: Channel Shuffle. DWConv: Depthwise Convolution.

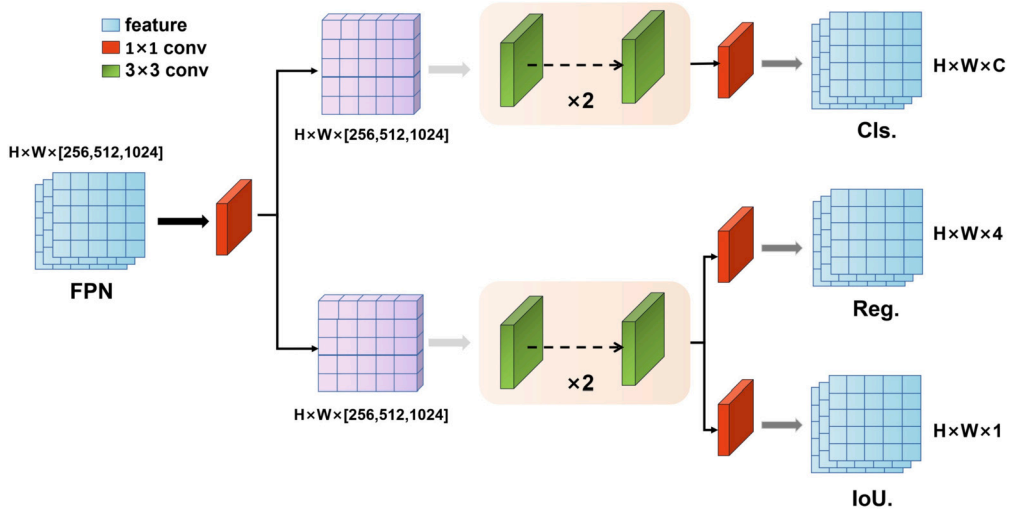

Fig. 6. Construction of Decoupled Head. Decoupled Head illustrates the independent pathways for classification and bounding box regression.

respectively. This integration emphasizes the method's efficiency and effectiveness in real-time object detection.

Oat diseases may present in diverse forms, including small spots and extensive areas of discoloration. The traditional Neck structure has limitations when fusing multi-scale features. The Slim-Neck module improves the multi-scale feature fusion capability through GSConv, enabling the model to adapt to the detection of diseased areas of different sizes and shapes, and enhancing the ability to recognize various disease forms.

2.3.3. Decoupled head

This decoupling mechanism, rarely explored in agricultural disease detection, enables more precise lesion localization in complex backgrounds, as evidenced by our improved attention heatmaps and detection accuracy. The Decoupled Head approach in object detection models separates the tasks of classification and bounding box regression into distinct sub-networks. This separation allows for independent optimization of each task, enhancing the performance and accuracy of the model (Fig. 6). The classification head predicts the class probabilities of detected objects. It processes the feature map generated by the backbone network through several layers optimized for distinguishing different object categories, resulting in a set of probability scores for each potential class. The formula for the classification head is as follows:

$$ClassProbability = Softmax(W_c \cdot F + b_c), \tag{4}$$

where W_c and b_c are the weights and biases for the classification head, and F is the input feature map.

The regression head predicts the coordinates of the bounding boxes for detected objects. It also processes the feature map from the backbone but focuses on spatial information to ensure accurate localization. The formula for the regression head is as follows:

$$Bounding\,Box Coordinates = W_r \cdot F + b_r, \tag{5}$$

where W_r and b_r are the weights and biases for the regression head, and F is the input feature map. The outputs from the classification and regression heads are combined to form the final detection results. Each detected object will have a bounding box from the regression head and an associated class probability from the classification head.

Intersection over Union (IoU) evaluates the overlap between the predicted bounding box and the true bounding box. It measures detection accuracy by calculating the ratio of the intersection area to the union area of the predicted and true boxes. A higher IoU indicates more accurate positioning [36].

By decoupling the tasks, the model can independently optimize the weights and architecture for each head, leading to better performance in both classification and localization. This approach minimizes conflicts between the objectives of classification and regression, allowing each head to focus on its task exclusively.

The distribution of disease symptoms on oat leaves can be extremely uneven, with lesions sometimes concentrated at the edges of the leaves and sometimes randomly scattered on the surface of the leaves. The Decoupled Head module separates the classification and localization tasks, allowing the model to independently optimize these two tasks, avoiding the loss of detection accuracy due to task conflicts. This design improves the model's performance when dealing with irregularly distributed disease symptoms, enabling it to more accurately detect complex disease areas.

3. Results and discussions

3.1. Experimental environment

To ensure consistent results, all subsequent experiments are conducted in the same experimental environment. The specific parameters are shown in Table 1. During training, the input size is set to 640×640 pixels, with a batch size of 16. The training process lasts for 300 epochs, with a patience parameter of 50. We train the model using Stochastic Gradient Descent, employing a 0.01 step size and a weight-decay coefficient of 5×10^{-4}. Data augmentation parameters include an HSV hue adjustment (hsv_h) of 0.015, an HSV saturation adjustment (hsv_s) of 0.7, an HSV value adjustment (hsv_v) of 0.4, a translation factor of 0.2, and a scaling factor of 0.9. The training is performed on a desktop platform equipped with an Intel(R) Xeon(R) Gold 6348 CPU @ 2.60GHz, 128 GB of RAM, and one NVIDIA A40 GPU with 40 GB of RAM, running on the Ubuntu 22.04.4 operating system. The runtime environment includes CUDA 12.2, cuDNN 8.9.7, PyTorch 2.3.1, and Python 3.8. Hyperparameters were first set following standard YOLO training practices and subsequently refined using a manual search strategy, where learning rate, batch size, and augmentation parameters were iteratively adjusted based on validation performance.

3.2. Evaluation criteria

To assess the performance of the RSD-YOLO for recognizing oat leaf disease severity, we use precision (P), recall (R), and mean average precision at 0.5 (mAP@0.5) as evaluation metrics. These metrics are widely used in the computer vision community for benchmarking detection tasks, and their definitions can be found in seminal works such as Everingham et al. [36] and Lin et al. [37].

Table 1
Hyperparameter and environment configuration for model training.

Parameter	Value
Input size	640×640 pixels
Batch size	16
Epochs	300
Early stopping patience	50
Optimizer	Stochastic Gradient Descent (SGD)
Learning rate (step size)	0.01
Weight decay	5×10^{-4}
Data Augmentation	
HSV hue adjustment (hsv_h)	0.015
HSV saturation adjustment (hsv_s)	0.7
HSV value adjustment (hsv_v)	0.4
Translation factor	0.2
Scaling factor	0.9
Hardware	
CPU	Intel Xeon Gold 6348 @ 2.60GHz
RAM	128 GB
GPU	NVIDIA A40 (40 GB)
Operating System	Ubuntu 22.04.4
Software Environment	
CUDA version	12.2
cuDNN version	8.9.7
PyTorch version	2.3.1
Python version	3.8

3.3. Oat leaf disease severity identification results

The identification results of oat leaf disease severity using the improved YOLOv7-tiny model are presented in Fig. 7. The model achieves high precision, recall, and mean mAP@0.5. Specifically, the improved YOLOv7-tiny, which integrates the ReXNet backbone, Slim-Neck modules, and Decoupled Head, attains a precision of 91.6%, recall of 90.8%, and mAP@0.5 of 88.5%. This performance surpasses the baseline YOLOv7-tiny model, which records a precision of 85.6%, recall of 80.8%, and mAP@0.5 of 80.9%. Additionally, experimental findings reveal that the model is adept at accurately detecting and classifying the disease severity of multiple oat leaves simultaneously. To provide further insight into the training process, Fig. 8 depicts the learning curves of the proposed RSD-YOLO model over 300 epochs. Both the training and validation loss decrease smoothly and converge stably, while training and validation precision steadily increase, with validation precision stabilizing at 91.6%. The close alignment between training and validation curves indicates that the model fits the data effectively without obvious overfitting, which further confirms the reliability of the reported results.

3.4. Comparative experiment

3.4.1. Different models performance comparison

To further demonstrate the superiority of the RSD-YOLO model in identifying oat leaf diseases, we compare various performance metrics precision, recall, mAP@0.5, and GFLOPs across different trained models on the test set. The RSD-YOLO is also evaluated against other models, including Faster R-CNN [38], DETR [39], RT-DETR [40], EfficientSAM [41], Internimage [42], SwiftFormer [43], YOLOv5 [44], YOLOv6 [45], YOLOv7, YOLOv8 [46], YOLOv9 [47], YOLOv10 [48], YOLOv11, and YOLOv12 [49].

The results in Table 2 demonstrate the superiority of RSD-YOLO. By integrating the Decoupled Head structure from YOLOv11 and optimizing the ReXNet Backbone and Slim-Neck, the RSD-YOLO model achieves 91.6% precision, 90.8% recall, and 88.5% mAP@0.5. These results greatly exceed other models in terms of identification accuracy, while the model requires only 6.5M parameters and 11.2 GFLOPs, making it lightweight and suitable for deployment on edge devices.

Compared to Faster R-CNN (42.5M parameters, 122.6 GFLOPs) and DETR (40.8M parameters, 90.2 GFLOPs), both of which rely on more

Fig. 7. Oat leaf disease severity identification results of RSD-YOLO: (a) Identification of H, R, and MR; (b) Identification of H and MS; (C) Identification of R and S.

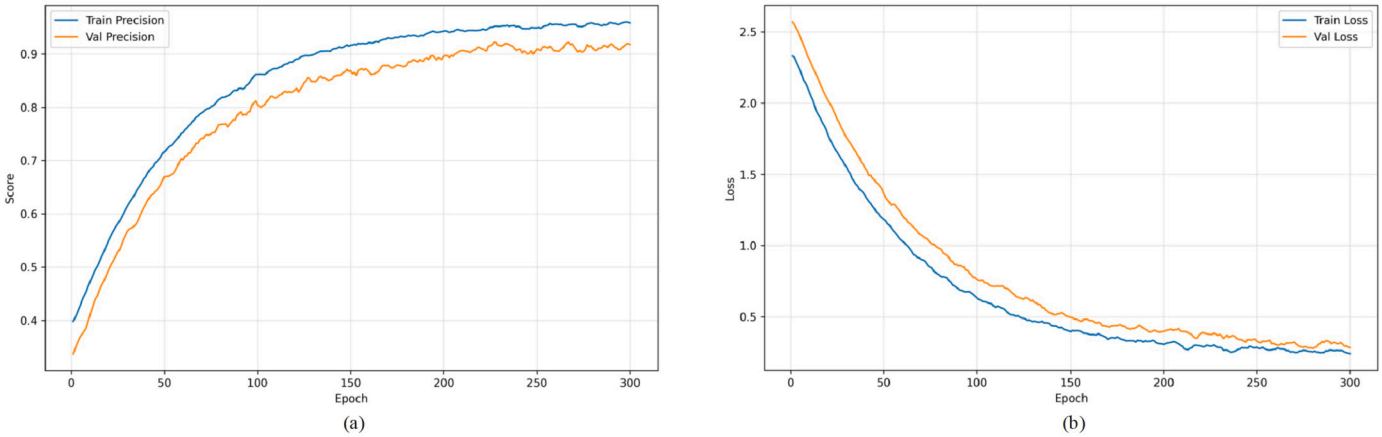

Fig. 8. Learning curves results of RSD-YOLO: (a) precision curves; (b) loss curves.

Table 2
Comparative experiment in oat leaf disease severity identification. The significance of bold means the best parameters and results.

Methods	Precision (%)	Recall (%)	mAP@0.5 (%)	Parameters (M)	GFLOPs ($\times 10^9$ FLOPs)
Faster R-CNN [38]	73.1	76.5	74.8	42.5	122.6
DETR [39]	81.6	82.2	81.3	40.8	90.2
Internimage [42]	88.2	85.1	86.3	50.3	73.2
SwiftFormer [43]	81.3	82.8	81.1	10.2	22.9
RT-DETR [40]	80.4	81.7	80.2	32.7	31.4
REfficientSAM [41]	82.7	81.9	80.5	13.6	16.8
YOLOv5s [44]	84.0	73.5	83.4	7.1	15.8
YOLOv5n [44]	74.1	80.5	81.4	**1.8**	**4.1**
YOLOv6 [45]	80.6	79.3	82.7	18.7	50.0
YOLOv7 [29]	88.8	86.0	82.6	36.9	103.2
YOLOv7-tiny [29]	85.6	80.8	80.9	6.2	13.2
YOLOv8s [46]	76.5	77.9	83.6	11.2	28.4
YOLOv8n [46]	74.2	83.5	85.1	3.2	8.1
YOLOv9s [47]	76.3	84.5	84.2	7.1	38.7
YOLOv10s [48]	81.1	79.9	81.7	7.2	24.5
YOLOv11s [30]	79.0	82.5	84.7	9.4	21.3
YOLOv12s [49]	81.5	83.4	82.5	9.3	25.3
RSD-YOLO (ours)	**91.6**	**90.8**	**88.5**	6.5	11.2

complex architectures, RSD-YOLO requires substantially fewer parameters and less computational cost, while delivering higher precision and recall. This highlights the efficiency and effectiveness of our design.

Compared to YOLOv5 and YOLOv6, which already exhibit good computational efficiency, RSD-YOLO achieves higher precision and recall with fewer resources. For example, YOLOv5s achieves a precision of 84.0% and a recall of 73.5% with 7.1M parameters, while YOLOv6 records 80.6% precision and 79.3% recall with 18.7M parameters and 50.0 GFLOPs. In contrast, RSD-YOLO achieves 91.6% precision and 90.8% recall with only 6.5M parameters and 11.2 GFLOPs. YOLOv5n, although extremely lightweight (1.8M parameters, 4.1 GFLOPs), performs poorly with only 74.1% precision, 80.5% recall, and 81.4% mAP@0.5, showing the clear trade-off between size and accuracy.

When compared to advanced models such as YOLOv12 and YOLOv11, RSD-YOLO maintains a competitive mAP@0.5 while reducing computational complexity and parameter count. For example, YOLOv11 achieves a mAP@0.5 of 84.7% with 9.4M parameters and 21.3 GFLOPs, whereas RSD-YOLO achieves a higher mAP@0.5 of 88.5% with fewer parameters (6.5M) and nearly half the computational cost. This balance of accuracy, efficiency, and compactness makes RSD-YOLO particularly

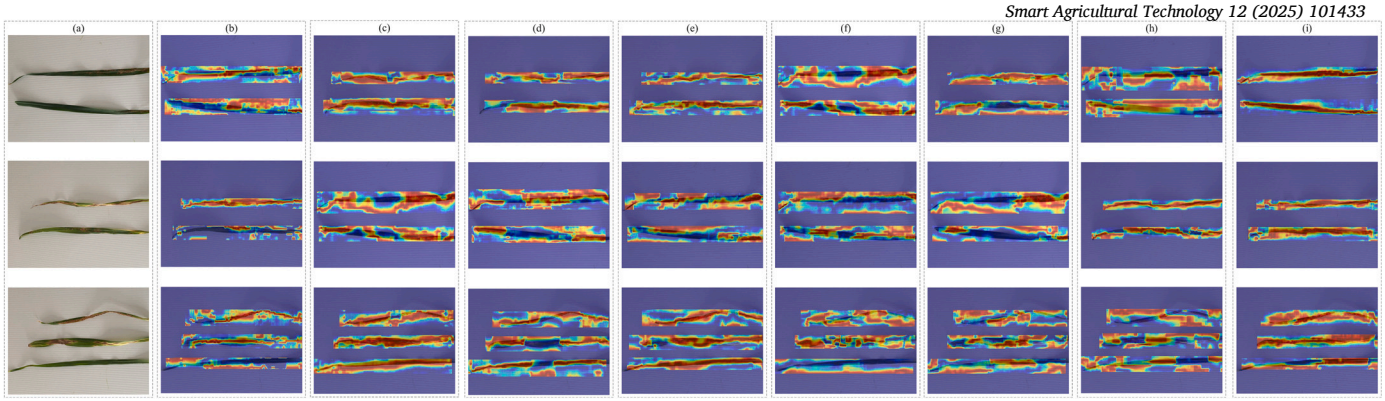

Fig. 9. Attention heat maps generated by different models: (a) Original images; (b) YOLOv5s; (c) YOLOv7-tiny; (d) YOLOv8s; (e) YOLOv9s; (f) YOLOv10s; (g) YOLOv11s; (h) YOLOv12s and (i) RSD-YOLO (ours).

suitable for identifying oat leaf disease severity on resource-constrained devices.

3.4.2. *Visual interpretability of the network*

XAI is a valuable tool for visually exploring hidden patterns within neural networks [50]. XAI helps users understand the inner workings of models by generating visual explanations of the model, providing details of the decision-making process, and revealing how much the model pays attention to different input features. In this study, we applied GradCAM XAI techniques to interpret the model's capabilities. Generally, more accurate algorithms exhibit higher consistency in visualizing attention. The heat maps are overlayed on the original images, emphasizing the predicted class regions in red (Fig. 9). As shown in the figure, the red area in the heat map generated by RSD-YOLO widely covers the entire leaf and is not limited to specific lesions. This attention method shows that RSD-YOLO not only relies on local features to judge diseases but also makes comprehensive judgments based on the characteristic information of the entire leaf, thus improving the robustness and accuracy of identification.

The comprehensive attention mechanism applied to the entire leaf offers considerable benefits for RSD-YOLO. In scenarios where disease distribution is irregular, the model can use information across various feature dimensions to detect potential disease characteristics, thereby minimizing the likelihood of oversight or misidentification resulting from variations in lesion presentation. This also reflects the model's sensitivity to subtle lesion characteristics in resource-constrained environments, allowing it to provide stable identification performance in practical agricultural scenarios.

3.5. *Ablation study*

3.5.1. *Analysis of the contribution of each module*

The present study adopts YOLOv7-tiny as the baseline model and incorporates ReXNet, Slim-Neck, and Decoupled Head modules to enhance its performance. The enhanced YOLOv7-tiny network was tested, with results in Table 3 and attention heat maps in Fig. 10. The baseline YOLOv7-tiny model achieves an accuracy of 85.6%, a recall of 80.8%, and a mAP@0.5 of 80.9%. By integrating the ReXNet backbone network, these metrics improved, with precision rising to 86.6%, recall to 84.7%, and mAP@0.5 to 81.1%. A further enhancement is observed after adding the Decoupled Head module, which increases precision to 88.4%, recall to 85.7%, and mAP@0.5 to 81.8%. The combined ReXNet, SlimNeck, and Decoupled Head model reached 91.6% precision, 90.8% recall, and 88.5% mAP@0.5, greatly enhancing performance.

1) Impact of ReXNet: After using ReXNet as the backbone network, the overall precision, recall, and mAP@0.5 of the model have all been improved. Experimental data shows that the introduction of ReXNet has increased accuracy from 85.6% to 86.6%, the recall rate from 80.8% to 84.7%, and the mAP@0.5 from 80.9% to 81.1%. Has increased accuracy from 85. 6% to 86. 6%, the recall rate from 80.8% to 84.7%, and the mAP@0.5 from 80.9% to 81.1%. ReXNet's integration greatly improved the model's feature extraction. As seen in Fig. 10, the attention heat map for the ReXNet-enhanced model shows a notable improvement in focusing on disease-affected regions. Compared to the baseline, the ReXNet model concentrates more effectively on subtle disease features, particularly those with low contrast. ReXNet improves feature extraction by optimizing channel dimensions between layers, achieving better identification of complex features with low computational complexity. Especially when dealing with low-contrast lesions, ReXNet demonstrates a stronger ability to capture details, thereby improving identification accuracy.

2) Impact of Slim-Neck: The addition of the Slim-Neck further optimizes the computational efficiency of the model. This module uses GSConv and VoV-GSCSP structures, which reduces computational overhead while ensuring the efficiency of the model. Based on this improvement, the accuracy of the model increases from 86.6% to 88.4%, the recall rate also increases from 84.7% to 85.7%, and the mAP@0.5 reaches 81.8%. Fig. 10 demonstrates that the Slim-Neck-enhanced model distributes its attention more evenly across disease regions of varying sizes, from small spots to elongated lesions. Slim-Neck combines standard convolution with DSC to reduce the parameters while maintaining rich feature expression capabilities. This is particularly effective for processing multi-scale features and improves the identification performance of the model in complex environments.

3) Impact of Decoupled Head: After the introduction of the Decoupled Head module, the model has greatly improved the model's classification and bounding box regression. This module separates classification and regression tasks, allowing each task to be optimized independently, reducing conflicts between tasks and improving identification accuracy and positioning accuracy. Experimental data shows that after integrating Decoupled Head, the precision reaches 91.6%, the recall increases to 90.8%, and mAP@0.5 finally reaches 88.5%. Fig. 10 shows that Decoupled Head allows the model to produce highly localized attention, reducing interference from irrelevant regions while enhancing the focus on disease lesions. The Decoupled Head enhances the model's ability to manage classification and positioning tasks in target identification. It greatly improves overall performance in assessing classification and positioning tasks in target identification. It greatly improves overall performance in assessing disease severity, making the model well-suited for real-time identification applications.

To further demonstrate the classification effectiveness of the proposed RSD-YOLO framework, we present confusion matrices for both the baseline YOLOv7-tiny model and RSD-YOLO, as shown in Fig. 11. These matrices provide a comprehensive view of class-wise predictions across the five disease severity categories. Compared to the baseline model, RSD-YOLO exhibits a clear improvement in classification accuracy across all categories. Notably, the proposed model reduces mis-

Table 3
Ablation experiment results. The significance of bold means the best parameters and results.

No.	ReXNet	Slim-Neck	Decoupled Head	Category	Precision (%)	Recall (%)	mAP@0.5 (%)	GFLOPs ($\times 10^9$ FLOPs)
1	-	-	-	H	89.0	78.5	82.1	13.2
				R	80.5	83.4	81.3	
				MR	86.2	79.1	79.6	
				MS	85.1	84	80.4	
				S	87.3	79	80.9	
				Average	85.6	80.8	80.9	
2	✓			H	89.5	85.3	80.5	11.3
				R	82.3	84.6	81.0	
				MR	87.1	84.5	82.3	
				MS	85.4	83.9	79.8	
				S	88.3	85.2	81.7	
				Average	86.5	84.7	81.1	
3		✓		H	87.3	86.2	87.5	10.8
				R	81.7	84	80.4	
				MR	85.9	87.1	81.2	
				MS	85.4	85.5	82.5	
				S	88.5	85.4	84.0	
				Average	85.6	85.6	83.1	
4			✓	H	87.4	90.2	87.2	14.2
				R	82.9	86.1	82.5	
				MR	85.5	88.9	83.1	
				MS	84.7	89.5	84.3	
				S	90.3	86.8	88.6	
				Average	86.2	88.3	85.1	
5	✓	✓		H	91.2	86.1	83	**10.4**
				R	82.9	84.2	82.5	
				MR	84.5	89	84.1	
				MS	85.4	84.7	82.7	
				S	89.7	85.6	86.3	
				Average	86.7	85.9	83.7	
6	✓		✓	H	90.1	87.3	84.2	12.1
				R	85.3	82.5	80.1	
				MR	87.9	86.7	81.7	
				MS	89.2	85.8	79.8	
				S	89.4	86.2	83.3	
				Average	88.4	85.7	81.8	
7		✓	✓	H	88.1	87.5	88.5	11.6
				R	82.4	82.2	84.2	
				MR	87.7	86.9	87.3	
				MS	87.3	84.7	84.6	
				S	89.1	90.3	89.0	
				Average	86.9	86.3	86.7	
8	✓	✓	✓	H	91.3	89.4	87.3	11.2
				R	89.6	91.2	88.7	
				MR	91.9	90.3	89.4	
				MS	92.4	92.7	90.1	
				S	92.8	90.4	86.9	
				Average	**91.6**	**90.8**	**88.5**	

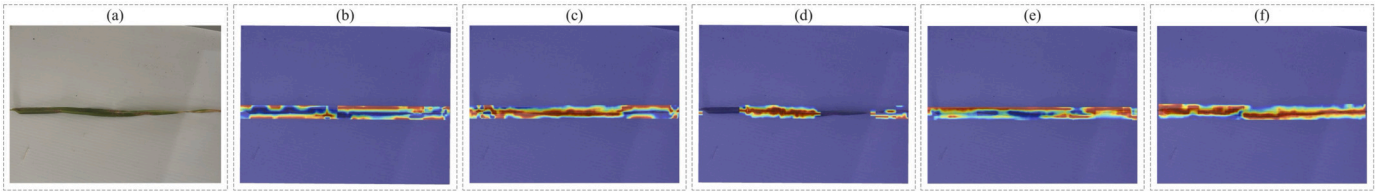

Fig. 10. Attention heat maps generated by different models: (a) Original image; (b) Baseline (YOLOv7-tiny); (c) Using ReXNet; (d) Using Slim-Neck; (e) Using Decoupled Head; and (f) RSD-YOLO (ours).

classification between adjacent classes such as MR and MS, which are often visually similar. For example, the true MR class achieves 705 correct predictions under RSD-YOLO, compared to 618 under the baseline. Similarly, the Healthy and Susceptible classes exhibit increased true positive rates and reduced misclassification into neighboring categories. Furthermore, the number of 'missed' or heavily misclassified samples decreases significantly in RSD-YOLO, as indicated by lighter cells in the off-diagonal regions.

Overall, the confusion matrix analysis reinforces the quantitative improvements shown in precision, recall, and mAP, and confirms that the proposed enhancements yield more reliable and fine-grained predictions of disease severity levels in oat leaves.

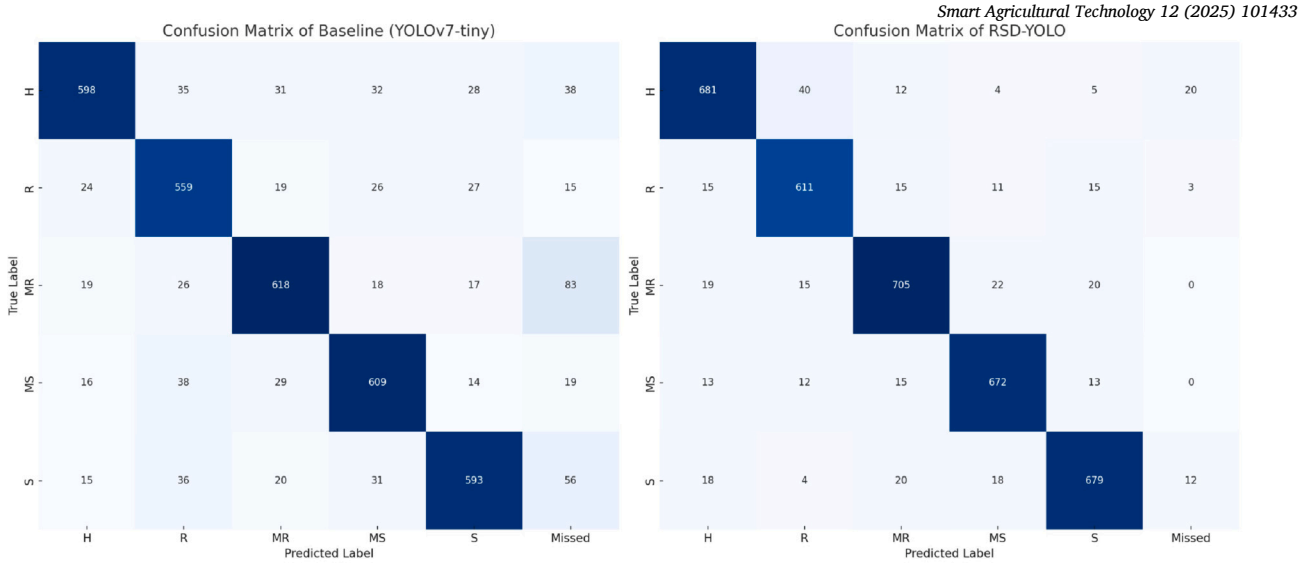

Fig. 11. Confusion Matrix of Baseline and RSD-YOLO.

Fig. 11. Confusion Matrix of Baseline and RSD-YOLO.

Table 4
Ablation experiment of different backbone architectures. The significance of bold means the best parameters and results.

Backbone	Precision (%)	Recall (%)	mAP@0.5 (%)	GFLOPS ($\times 10^9$ FLOPs)
Baseline (yolov7-tiny) [29]	85.6	80.8	80.9	13.2
MobileNetV3 [51]	84.9	79.2	80.2	**9.8**
ResNet50 [52]	86.2	84.3	**81.9**	15.6
EfficientNet [53]	85.8	82.6	80.9	12.4
ReXNet [31]	**86.6**	**84.7**	81.1	11.3

3.5.2. Ablation experiment of different backbones

To further demonstrate the benefits of using ReXNet as the backbone, a specialized ablation study is conducted, focusing solely on different backbone architectures while keeping the other components of YOLOv7-tiny unchanged. The objective is to demonstrate the effectiveness of ReXNet in identifying the severity of oat leaf diseases through a comparative analysis with other leading backbone models. This analysis also examines the unique characteristics of ReXNet and oat leaf images to explain why ReXNet is well-suited for this task. Various state-of-the-art backbones are tested, including MobileNetV3 [51], ResNet50 [52], EfficientNet [53], and ReXNet. The neck and head structures of YOLOv7-tiny remain unchanged during these experiments. Performance metrics for comparison include precision, recall, mAP@0.5, and GFLOPS. As shown in Table 4, ReXNet outperforms other mainstream backbones in oat leaf disease identification by achieving the best balance between performance and inference speed. ReXNet excels by effectively capturing complex features in oat leaf images, essential for accurate disease detection. Its design enables precise identification while maintaining relatively low computational complexity. While other backbones are effective, none achieve ReXNet's precision and recall, emphasizing its fit for this application.

ReXNet employs a linear parameterization of channel dimensions across layers, ensuring a gradual and consistent increase in the model's capacity to capture features. This design is particularly adept at capturing subtle variations and complex patterns in oat leaf images, which are essential for precise disease identification. What's more, ReXNet uses DSC and expanded layers, enabling the model to preserve a simplified architecture while effectively capturing features across multiple scales. This capability is essential for handling high-resolution oat leaf images where disease symptoms can vary in scale and intensity.

Oat leaf images often exhibit complex textures and low-contrast features, making disease identification challenging. ReXNet's versatile channel configuration and enhanced feature layers are particularly effective in capturing and emphasizing these critical details, resulting in improved accuracy for identifying and classifying disease severity. ReXNet's ability to balance computational efficiency with advanced feature extraction makes it well-suited for oat leaf disease identification. The ablation study confirms that ReXNet provides superior performance compared to other backbones, making it the optimal choice for this specific application.

3.6. Discussion

The proposed RSD-YOLO framework exhibits significant potential in the task of oat leaf disease severity identification. In this section, we summarize its key advantages and current limitations based on experimental results and architectural analysis.

3.6.1. Advantages

The most prominent advantage of RSD-YOLO lies in its remarkable detection performance under constrained computational resources. The proposed model integrates ReXNet, Slim-Neck, and Decoupled Head modules within a unified framework. ReXNet improves the network's ability to extract fine-grained features from low-contrast lesion regions by optimizing channel configurations and employing DSC. The Slim-Neck enhances the fusion of multi-scale features, which is crucial for accurately detecting disease symptoms of varying sizes and shapes. Meanwhile, the Decoupled Head separates classification and localization tasks, reducing task interference and enabling more stable predictions, especially when lesions overlap or appear along complex leaf structures. As a result of these targeted design choices, the model

achieves a precision of 91.6%, recall of 90.8%, and mAP@0.5 of 88.5%. Compared to the baseline YOLOv7-tiny (85.6% precision, 80.8% recall, 80.9% mAP@0.5), this represents significant improvements in accuracy.

In addition, RSD-YOLO maintains a lightweight and efficient design, achieving only 11.2 GFLOPs while outperforming larger and more computationally expensive models such as Faster R-CNN and DETR. This characteristic enables real-time inference and edge deployment on resource-constrained platforms, such as drones, or mobile agricultural devices—making it highly practical for on-farm disease monitoring.

Furthermore, YOLOv7-tiny was chosen as the baseline model for improvement due to its excellent trade-off between detection accuracy and inference speed among lightweight object detectors. Compared to other tiny variants such as YOLOv5n, YOLOv6, or YOLOv8n, YOLOv7-tiny demonstrates higher baseline precision (85. 6% precision and 80. 9% mAP @ 0.5) with moderate computational complexity (13.2 GFLOPs). Its modular structure also allows easier integration and replacement of components such as the backbone and head. These characteristics make YOLOv7-tiny a more promising and stable foundation for real-time applications in precision agriculture, especially when further optimized with advanced modules like ReXNet, Slim-Neck, and Decoupled Head.

Another important merit of our design lies in the modular nature of its components. Each module contributes independently to the model's performance, as verified through ablation experiments. This not only confirms the effectiveness of each innovation but also allows the framework to be flexible and extensible for future improvements or cross-domain applications.

Lastly, the constructed dataset with five expert-annotated severity categories and comprehensive data augmentation further supports the robustness and generalization of the model. While the current dataset is oat-specific, the framework itself is general and can be adapted to other crops with similar annotation protocols.

3.6.2. Limitations

Despite its advantages, the current framework presents a few limitations. First, the model's effectiveness remains dependent on the quality and variability of the dataset. Although the imaging conditions are controlled during data collection, such consistency is difficult to guarantee in real-world fields, where variations in lighting, occlusion, leaf overlap, and environmental noise may degrade performance.

Second, although the model is designed to be lightweight, real-time large-scale deployment across entire fields may still require further optimization. Techniques such as model pruning, quantization, or hardware-specific acceleration will be needed to enhance processing speed and reduce memory consumption for seamless operation on embedded platforms.

Lastly, the model currently relies on RGB images. While these are accessible and practical, integrating additional sensing modalities such as hyperspectral or multispectral imaging could further improve early-stage disease identification and expand the system's application in more complex or mixed infection scenarios.

4. Conclusions

In this study, RSD-YOLO, a novel and lightweight framework for real-time identification of oat leaf disease severity, has been presented. The model incorporates three innovative modules: a Regularized Xception-based backbone (ReXNet) for efficient and discriminative feature extraction, a Slim-Neck module for multi-scale fusion with low computational cost, and a Decoupled Head separates classification and regression for improved precision and localization. Extensive experiments results have demonstrated that RSD-YOLO achieves superior performance compared to existing YOLO variants and mainstream detectors, with 91.6% precision, 90.8% recall, and 88.5% mAP@0.5, all while maintaining a lightweight design of 11.2 GFLOPs. Ablation studies have also confirmed that each module contributes meaningfully to the final performance, underscoring the robustness and efficiency of the overall architecture.

The RSD-YOLO addresses the key challenges in crop disease monitoring, such as low-contrast lesions, complex leaf textures, and deployment under limited computational resources. These attributes make RSD-YOLO particularly suitable for edge-based precision agriculture systems, including unmanned aerial vehicles, mobile platforms, and smart farming devices.

In future work, we plan to: (i) Expand the dataset to include more oat cultivars and real-field images under varying conditions, (ii) Generalize the model to other crops and diseases by transfer learning, (iii) Explore the integration of hyperspectral and temporal information for early disease detection, (iv) Apply model compression or neural architecture search (NAS) to further enhance deployment efficiency. We believe this study offers a practical, scalable, and accurate solution for sustainable crop monitoring, and provides a foundation for further innovations in smart agriculture.

CRediT authorship contribution statement

Yongquan Zhang: Methodology, Funding acquisition, Conceptualization. **Yiwei Xu:** Writing – review & editing, Writing – original draft, Software, Formal analysis. **Taosheng Xu:** Writing – review & editing. **Changmiao Wang:** Writing – review & editing. **Chengdao Li:** Resources, Data curation, Conceptualization. **Hai Wang:** Writing – review & editing, Resources, Project administration, Conceptualization.

Declaration of competing interest

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

Acknowledgements

The disease nursery was established by Dr Manisha Shankar, Dr Hieu Tran and Mr Ethan Ryan in the projects funded by Grains Research and Development Corporation (UMU2404-010RTX) and Western Australia Oat Industry Partnership (POIGP/03). We acknowledge the partial financial support provided by China Ministry of Education Humanities and Social Sciences Research-Planning Fund Project (No. 23YJA910007) and Zhejiang University of Finance and Economics. We also acknowledge the strategic and in-kind support by Murdoch University Harry Butler Institute.

Data availability

Data will be made available on request.

References

[1] Vera Rajičić, Vera Popović, Dragan Terzić, Dragan Grčak, Marijana Dugalić, Mihailović, et al., Impact of lime and npk fertilizers on yield and quality of oats on pseudogley soil and their valorisation, Not. Bot. Horti Agrobot. Cluj-Napoca 48 (4) (2020) 2134–2152.

[2] Sangami Ravindran, Probiotic oats milk drink with microencapsulated lactobacillus plantarum–an alternative to dairy products, Nutr. Food Sci. 51 (3) (2021) 471–482.

[3] Amandeep Sandhu, S.S. Dhaliwal, A.K. Shukla, Vivek Sharma, Ravinder Singh, Fodder quality improvement and enrichment of oats with cu through biofortification: a technique to reduce animal malnutrition, J. Plant Nutr. 43 (10) (2020) 1378–1389.

[4] FAO, FAOSTAT: crops and livestock products visualization, 2024. (Accessed 30 October 2024).

[5] B.D. Van Niekerk, Z.A. Pretorius, W.H.P. Boshoff, Potential yield losses caused by barley leaf rust and oat leaf and stem rust to South African barley and oat cultivars, S. Afr. J. Plant Soil 18 (3) (2001) 108–113.

[6] Clive H. Bock, Kuo-Szu Chiang, Emerson M. Del Ponte, Plant disease severity estimated visually: a century of research, best practices, and opportunities for improving methods and practices to maximize accuracy, Trop. Plant Pathol. 47 (1) (2022) 25–42.

[7] Yimy E. García-Vera, Andrés Polochè-Arango, Camilo A. Mendivelso-Fajardo, Félix J. Gutiérrez-Bernal, Hyperspectral image analysis and machine learning techniques for crop disease detection and identification: a review, Sustainability 16 (14) (2024) 6064.

[8] Deepti Barhate, Sunil Pathak, Bhupesh Kumar Singh, Amit Jain, Ashutosh Kumar Dubey, A systematic review of machine learning and deep learning approaches in plant species detection, Smart Agric. Technol. (2024) 100605.
[9] Yifan Bai, Junzhen Yu, Shuqin Yang, Jifeng Ning, An improved yolo algorithm for detecting flowers and fruits on strawberry seedlings, Biosyst. Eng. 237 (2024) 1–12.
[10] Seyed Mohamad Javidan, Ahmad Banakar, Kamran Rahnama, Keyvan Asefpour Vakilian, Yiannis Ampatzidis, Feature engineering to identify plant diseases using image processing and artificial intelligence: a comprehensive review, Smart Agric. Technol. (2024) 100480.
[11] Imane Bouacida, Brahim Farou, Lynda Djakhdjakha, Hamid Seridi, Muhammet Kurulay, Innovative deep learning approach for cross-crop plant disease detection: a generalized method for identifying unhealthy leaves, Inf. Process. Agric. 12 (1) (2025) 54–67.
[12] Fu Zhang, Ruofei Bao, Baoping Yan, Mengyao Wang, Yakun Zhang, Sanling Fu, Lsannet: a lightweight convolutional neural network for maize leaf disease identification, Biosyst. Eng. 248 (2024) 97–107.
[13] M.P. Aishwarya, Padmanabha Reddy, Ensemble of cnn models for classification of groundnut plant leaf disease detection, Smart Agric. Technol. 6 (2023) 100362.
[14] Abhishek Upadhyay, Abhishek Patel, Abhishek Patel, Narendra Singh Chandel, Subir Kumar Chakraborty, Dattatray G. Bhalekar, Leveraging ai and ml in precision farming for pest and disease management: benefits, challenges, and future prospects, in: Ecologically Mediated Development: Promoting Biodiversity Conservation and Food Security, 2025, pp. 511–528.
[15] Md. Jawadul Karim, Md. Omaer Faruq Goni, Md. Nahiduzzaman, Mominul Ahsan, Julfikar Haider, Marcin Kowalski, Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight cnn architecture and grad-cam, Sci. Rep. 14 (2024) 16022.
[16] Shruthi B S, Narasimha Murthy M S, Chidurala Sai Prakash, Ramy Riad Al–Fatlawy, S. Meenakshi Sundaram, Plant leaf disease detection and classification using mobilenetv3 and wavelet kernel extreme learning machine, in: 2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), 2024, pp. 01–06.
[17] Tej Bahadur Shahi, Cheng-Yuan Xu, Arjun Neupane, William Guo, Recent advances in crop disease detection using uav and deep learning techniques, Remote Sens. 15 (9) (2023).
[18] Yang Lyu, Xiongzhe Han, Pingan Wang, Jae-Yeong Shin, Min-Woong Ju, Unmanned aerial vehicle-based rgb imaging and lightweight deep learning for downy mildew detection in kimchi cabbage, Remote Sens. 17 (14) (2025).
[19] Guan Wang, Yu Sun, Jianxin Wang, Automatic image-based plant disease severity estimation using deep learning, Comput. Intell. Neurosci. 2017 (1) (2017) 2917536.
[20] Wenhao Su, Jiajing Zhang, Ce Yang, Rae Page, Tamas Szinyei, Cory D. Hirsch, et al., Automatic evaluation of wheat resistance to Fusarium head blight using dual mask-rcnn deep learning frameworks in computer vision, Remote Sens. 13 (1) (2020) 26.
[21] Qiaokang Liang, Shao Xiang, Yucheng Hu, Gianmarc Coppola, Dan Zhang, Wei Sun, Pd2se-net: computer-assisted plant disease diagnosis and severity estimation network, Comput. Electron. Agric. 157 (2019) 518–529.
[22] Shuo Chen, Kefei Zhang, Yindi Zhao, Yaqin Sun, Wei Ban, Yu Chen, et al., An approach for rice bacterial leaf streak disease segmentation and disease severity estimation, Agriculture 11 (5) (2021) 420.
[23] Junfeng Gao, Jesper Cairo Westergaard, Ea Høegh Riis Sundmark, Merethe Bagge, Erland Liljeroth, Erik Alexandersson, Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning, Knowl.-Based Syst. 214 (2021) 106723.
[24] José G.M. Esgario, Renato A. Krohling, José A. Ventura, Deep learning for classification and severity estimation of coffee leaf biotic stress, Comput. Electron. Agric. 169 (2020) 105162.
[25] Gianni Fenu, Francesca Maridina Malloci, Using multioutput learning to diagnose plant disease and stress severity, Complexity 2021 (1) (2021) 6663442.
[26] Maheswari Prabhakar, Raja Purushothaman, Durga Prasad Awasthi, Deep learning based assessment of disease severity for early blight in tomato crop, Multimed. Tools Appl. 79 (2020) 28773–28784.
[27] Qingmao Zeng, Xinhui Ma, Baoping Cheng, Erxun Zhou, Wei Pang, Gans-based data augmentation for citrus disease severity detection using deep learning, IEEE Access 8 (2020) 172882–172891.
[28] Juliano P. Goncalves, Francisco A.C. Pinto, Daniel M. Queiroz, Flora M.M. Villar, Jayme G.A. Barbedo, Emerson M. Del Ponte, Deep learning architectures for semantic segmentation and automatic estimation of severity of foliar symptoms caused by diseases or pests, Biosyst. Eng. 210 (2021) 129–142.
[29] Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan Mark Liao, Yolov7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 7464–7475.
[30] Glenn Jocher, Jing Qiu, Ultralytics yolo11, 2024.
[31] Dongyoon Han, Sangdoo Yun, Byeongho Heo, YoungJoon Yoo, Rethinking channel dimensions for efficient model design, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 732–741.
[32] Hulin Li, Jun Li, Hanbing Wei, Zheng Liu, Zhenfei Zhan, Qiliang Ren, Slim-neck by gsconv: a lightweight-design for real-time detector architectures, J. Real-Time Image Process. 21 (3) (2024) 62.
[33] Maxim Tkachenko, Mikhail Malyuk, Andrey Holmanyuk, Nikolai Liubimov, Label studio: data labeling software, 2020-2025, Open source software available from https://github.com/HumanSignal/label-studio.
[34] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, Yolov4: optimal speed and accuracy of object detection, arXiv preprint, arXiv:2004.10934, 2020.
[35] Alexander Buslaev, Vladimir I. Iglovikov, Eugene Khvedchenya, Alex Parinov, Mikhail Druzhinin, Alexandr A. Kalinin, Albumentations: fast and flexible image augmentations, Information 11 (2) (2020).
[36] Mark Everingham, Luc Van Gool, Christopher K.I. Williams, John Winn, Andrew Zisserman, The Pascal visual object classes (voc) challenge, Int. J. Comput. Vis. 88 (2) (2010) 303–338.
[37] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, C. Lawrence Zitnick, Microsoft coco: common objects in context, in: European Conference on Computer Vision, Springer, 2014, pp. 740–755.
[38] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, Faster r-cnn: towards real-time object detection with region proposal networks, IEEE Trans. Pattern Anal. Mach. Intell. 39 (6) (2016) 1137–1149.
[39] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, Sergey Zagoruyko, End-to-end object detection with transformers, in: European Conference on Computer Vision, Springer, 2020, pp. 213–229.
[40] Yian Zhao, Wenyu Lv, Shangliang Xu, Jinman Wei, Guanzhong Wang, Qingqing Dang, et al., Detrs beat yolos on real-time object detection, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 16965–16974.
[41] Yunyang Xiong, Bala Varadarajan, Lemeng Wu, Xiaoyu Xiang, Fanyi Xiao, Chenchen Zhu, Xiaoliang Dai, Dilin Wang, Fei Sun, Forrest Iandola, et al., Efficientsam: leveraged masked image pretraining for efficient segment anything, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 16111–16121.
[42] Wenhai Wang, Jifeng Dai, Zhe Chen, Zhenhang Huang, Zhiqi Li, Xizhou Zhu, et al., Internimage: exploring large-scale vision foundation models with deformable convolutions, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 14408–14419.
[43] Abdelrahman Shaker, Muhammad Maaz, Hanoona Rasheed, Salman Khan, Ming-Hsuan Yang, Fahad Shahbaz Khan, Swiftformer: efficient additive attention for transformer-based real-time mobile vision applications, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 17425–17436.
[44] Glenn Jocher, Ultralytics yolov5, 2020.
[45] Chuyi Li, Lulu Li, Hongliang Jiang, Kaiheng Weng, Yifei Geng, Liang Li, et al., Yolov6: a single-stage object detection framework for industrial applications, arXiv preprint, arXiv:2209.02976, 2022.
[46] Rejin Varghese, M. Sambath, Yolov8: a novel object detection algorithm with enhanced performance and robustness, in: 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems, 2024, pp. 1–6.
[47] Chien-Yao Wang, I-Hau Yeh, Hong-Yuan Mark Liao, Yolov9: Learning what you want to learn using programmable gradient information, 2024, pp. 1–21.
[48] Ao Wang, Hui Chen, Lihao Liu, Kai Chen, Zijia Lin, Jungong Han, et al., Yolov10: real-time end-to-end object detection, Adv. Neural Inf. Process. Syst. 37 (2024) 107984–108011.
[49] Yunjie Tian, Qixiang Ye, David Doermann, Yolov12: attention-centric real-time object detectors.
[50] David Gunning, Mark Stefik, Jaesik Choi, Timothy Miller, Simone Stumpf, Guangzhong Yang, Xai—explainable artificial intelligence, Sci. Robot. 4 (37) (2019) eaay7120.
[51] Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, et al., Searching for mobilenetv3, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 1314–1324.
[52] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[53] Mingxing Tan, Quoc Le, Efficientnet: Rethinking model scaling for convolutional neural networks, 2019, pp. 6105–6114.