YOLOv8-RD: High-Robust Pine Wilt Disease Detection Method Based on Residual Fuzzy YOLOv8

Junchao Yuan, Lina Wang, Tingting Wang, Ali Kashif Bashir, Maryam M. Al Dabel, Jiaxing Wang, Hailin Feng, Kai Fang, and Wei Wang

Abstract—Pine Wilt Disease (PWD) poses a severe threat to the health of pine trees and has resulted in substantial losses to global pine forest resources. Due to the minute size of the pathogens and the concealed symptoms of PWD, early detection through remote sensing image technology is essential. However, in practical applications, remote sensing images are easily affected by factors such as cloud cover and changes in illumination, resulting in significant noise and blurriness in the images. These interference factors significantly reduce the accuracy of existing object detection models. Therefore, this paper presents a novel and highly robust methodology for detecting PWD, termed YOLOv8-RD. We synthesized the benefits of residual learning and Fuzzy Deep Neural Networks (FDNN) to develop a Residual Fuzzy module (ResFuzzy), which adeptly filters image noise and refines background features with enhanced smoothness. Simultaneously, we integrated a Detail Processing Module (DPM) into the ResFuzzy module to enhance the low-frequency detail features transmitted in residual learning. Furthermore, by incorporating the Dynamic upSampling operator (DySample), our model can dynamically adjust the sampling step size based on the variations in the input feature map during the upsampling process, thereby effectively recovering detail from the feature map. Our model exhibited exceptional robustness to severe noise. When evaluated on a PWD dataset with 100% interference samples at an intensity of 0.07, our model achieved an average precision improvement of 4.9%, 6.3%, 7.3%, and 3.0% compared to four most representative models, making it well-suited for PWD detection in interfering environments.

Index Terms—Remote Sensing, Interference Environments, Residual Learning, Fuzzy Deep Neural Networks, Upsampling Process.

Junchao Yuan, Lina Wang, Hailin Feng, and Kai Fang are with the College of Mathematics and Computer Science, Zhejiang A & F University, Hangzhou 311300, China. (Email: yjc955855@gmail.com, Linawangzafu@gmail.com, hlfeng@zafu.edu.cn, and Kaifang@ieee.org)

Tingting Wang is with the College of Computer Science and Engineering, Faculty of Innovation Engineering, Macau University of Science and Technology, Macau SAR 999078, China. (Email: tingtingwang@ieee.org)

Ali Kashif Bashir is with the Department of Computing and Mathematics, Manchester Metropolitan University, UK; Woxsen School of Business, Woxsen University, Hyderabad, India 502 345; Department of Computer Science and Mathematics, Lebanese American University, Beirut, Lebanon. (Email: dr.alikashif.b@ieee.org)

Maryam M. Al Dabel is with the Department of Computer Science and Engineering, University of Hafr AlBatin, Saudi Arabia (Email: maldabel@uhb.edu.sa)

Jiaxing Wang is with the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310014, China (Email: wjx@zjut.edu.cn)

Wei Wang is with the Guangdong-Hong Kong-Macao Joint Laboratory for Emotional Intelligence and Pervasive Computing, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, 518172, Guangdong, China (Email: ehomewang@ieee.org)

Corresponding authors: Hailin Feng and Kai Fang

I. INTRODUCTION

AS an important economic and ecological resource, pine trees play a crucial role in maintaining the stability of forest ecosystems, ensuring a steady supply of wood, sequestering carbon, and releasing oxygen [1]. However, PWD as one of the most destructive forestry diseases worldwide, poses a serious threat to pine resources. Its rapid spread and difficulty in early detection often lead to extensive damage to pine forests, causing severe economic losses and ecological consequences [2, 3]. Advancements in machine vision technology have led to the increasingly mature application of remote sensing image object detection in forestry monitoring of pests and diseases, particularly in the early warning and precise positioning of PWD [4].

The existing object detection models primarily focus on addressing issues such as occlusion, low resolution, and complex backgrounds [5]. Wang X et al. [6] optimized the YOLOv8 model to improve the detection ability of greenhouse vegetable pests and diseases under occlusion conditions. They incorporated an Occlusion-aware Attention Module (OAM) to endow the model with advanced capabilities for the precise detection of partially obscured or occluded objects. Additionally, a small object detection layer and HIOU loss function were introduced to improve the overall detection accuracy of vegetable diseases. Tian Y et al. [7] proposed a multi-scale dense detection method, named MD-YOLO, for small lepidopteran pests on sticky traps. This method utilizes an Adaptive Attention Module (AAM) to enhance the model's attention to image details. Moreover, the feature fusion process was optimized through the integration of multi-scale feature maps, thereby augmenting the model's efficacy in detecting objects across a range of scales.

However, during the process of remote sensing image acquisition by Unmanned Aerial Vehicle (UAV), atmospheric scattering, cloud cover, and changes in illumination can introduce noise into the image data [8]. This noise can manifest as random pixel-level fluctuations or larger-scale image blurriness and distortion, all of which significantly degrading image quality. Existing object detection models often suffer from low accuracy, false positives, and missed detection in noisy conditions. This not only hinders forestry management personnel from taking timely and effective control measures but also risks missing the window of optimal prevention due to misjudgment, further exacerbating the spread of PWD and

causing severe damage to the growth of pine trees.

Therefore, developing anti-interference detection algorithms for detecting PWD in remote sensing images and improving the accuracy of object detection models in interfered environments is of profound significance for the sustainable development of modern forestry. Such advancements not only ensure the rational use and effective protection of forestry resources and maintain biodiversity, but also enhance the ability to predict pest and disease outbreaks, reduce economic losses caused by these issues, and ensure the stability and balance of ecosystems.

YOLO algorithms, renowned for their efficient detection speed and powerful feature extraction capabilities, are commonly employed in forest pest monitoring tasks. In this study, we would utilize the YOLOv8 algorithm for pine wilt disease detection in UAV remote sensing images.

To enhance the model's robustness to noise, this study integrates FDNN into the YOLOv8 model. FDNN combine the learning capabilities of neural networks with the noise handling capabilities of fuzzy logic [9]. By fuzzifying the input data, the impact of noise on the model's prediction results is reduced, thereby enhancing the model's robustness to noisy images [10]. Meanwhile, FDNN offer a promising approach to plant disease and pest detection. In plant disease and pest detection, the background features are often complex, while disease features are relatively small. Therefore, FDNN can be employed to blur the complex background features and effectively enhance the disease features [11]. Moreover, the symptoms of plant diseases and pests often vary depending on the plant species and the stage of disease development. FDNN can handle these fuzzy features, enabling them to make effective judgments even when faced with different disease symptoms [12]. To extract complex features of tomato diseases and pests, Tian X et al. [13] improved the fuzzy inference layer and fuzzy pooling layer of FDNN, enhancing the model's detection accuracy. Koshariya A K et al. [14] analyzed the application of FDNN in plant disease and pest detection and risk assessment. By optimizing fuzzy rules and fuzzification processing, they improved feature extraction in complex environments.

The YOLOv8n model was selected as the base network in this study and has been modified to fulfill the requirements for anti-interference detection in remote sensing images. The main contributions of this paper are as follows:

(1) To tackle the challenges posed by cloud cover and illumination variability in PWD images, we propose an innovative ResFuzzy module that synergistically integrates the advantages of residual learning with fuzzy deep neural networks. The module effectively filters image noise and suppresses background features through multiple residual blocks and fuzzy layers. Moreover, we integrate a DPM module within the ResFuzzy framework to significantly bolster the model's capacity for detecting and discerning small objects in remote sensing images.

(2) We improved the upsampling process in YOLOv8n by introducing the DySample module to mitigate detail loss caused by noise. This module equips the model with the capability to dynamically modulate the sampling step size in response to variations in input features, thereby refining the sensitivity and adaptability of the sampling process to fluctuating input characteristics. Consequently, the robustness of the object detection model under noisy conditions is boosted.

(3) Our proposed YOLOv8-RD model effectively addresses the challenges of cloud occlusion and lighting variations in detecting PWD, thereby transcending the performance constraints of conventional detection models. The YOLOv8-RD model not only performs well on non-interference datasets but also demonstrates high robustness under various interference conditions. The employment of this model facilitates precise detection of infected PWD areas in remote sensing imagery, even under challenging conditions, thereby empowering forestry management personnel to execute timely interventions and substantially mitigate economic losses attributable to the disease.

II. RELATED WORK

A. Disease and Pest Detection

The existence of diseases and pests severely harms the plant health. Liu B et al. [15] proposed DAC-PPYOLOE to promote the detection accuracy of apple pests in complex environments. This model utilizes an adaptive feature fusion strategy with residual connectivity and deep separable convolution to effectively leverage deep and shallow feature maps for small object detection. Qi J et al. [16] introduced the improved SE-YOLOv5 network for tomato diseases and pest detection. The SE attention mechanism enhances key feature extraction, overcoming the limitations of existing methods in feature screening and model generalization. Deng F et al. [17] proposed a Federated Learning (FL) based Faster R-CNN model to tackle issues concerning data imbalance, diversity, and complex detection environments in traditional plant disease and pest detection. This model utilizes FL's distributed computing to reduce data storage and communication costs. Additionally, ResNet-101 replaces VGG-16 in the convolutional layer, improving multi-scale detection accuracy for various diseases and pests. Irianto S Y et al. [18] proposed a corn leaf disease detection method based on Fuzzy C-means (FCM) and Long Short-Term Memory (LSTM) algorithms to improve the detection accuracy of corn diseases. After extracting texture features from disease images using Gray Level Co-occurrence Matrix (GLCM), they were fed into the LSTM algorithm for classification and achieved an accuracy of 80.24%. Chang B et al. [19] addressed the challenges of complex plant disease features and limited datasets by proposing an Edge Feature Guidance (EFG) module to enhance the model's ability to extract local edge features. The EFG module can be integrated into vision transformers like ViT and Swin, enabling the model to incorporate multi-scale features and edge information, thereby improving its overall performance.

B. Small object Detection

Small objects in remote sensing imagery, owing to their constrained pixel coverage and attenuated feature intensity relative to both the background and larger objects, are exceptionally vulnerable to being obscured by background noise. Hou Q et

al. [20] proposed RISTDnet, a deep learning-based network for infrared small object detection. This network addresses low image contrast and low signal-to-noise ratios in complex backgrounds. It combines manual feature methods with convolutional neural networks for feature extraction and utilizes likelihood graph thresholds for real object segmentation. Experiments demonstrate the network's ability to accurately detect small objects. Dai Y et al. [21] proposed a network model for infrared small object detection. The model leverages a deep, parameter-free, non-linear feature refinement layer specially designed to extract long-range dependencies between features. Additionally, it employs bottom-up attention modulation to integrate low-level details into higher-level features, effectively preserving small object information. Duan H et al. [22] designed an adaptive mechanism algorithm inspired by the physiological characteristics of eagle vision. Based on the physiological structure of eagle vision, this algorithm establishes a mathematical model capable of adapting to various environmental interferences, thereby enhancing the model's ability to detect small objects in complex and variable marine environments.

C. Interference Environment Detection

In various object detection tasks, image quality may suffer due to inclement weather conditions, leading to reduced model accuracy. This, in turn, can negatively impact production, safety, and daily activities. To address the challenge of low detection accuracy in low-quality images captured under severe weather conditions, Liu W et al. [23] proposed an adaptive object detection framework IA-YOLO. They introduced the DIP of the adjustable image processing module in the small convolutional neural network to enhance image quality and improve object detection accuracy. To confront the dilemma of severe weather detection, Qin et al. [24] proposed DEnet, a detection-driven network comprised of three key modules. Splitting images into low and high-frequency components using the Laplace pyramid and forming global enhancement components with various convolution kernels significantly boost detection accuracy. Cui Z et al. [25] introduced the Multi-task Automatic Encoding Transformation (MAET) model for nighttime object detection, effectively enhancing detection accuracy in low-light conditions.

Existing object detection techniques have made notable strides in plant disease detection, small object detection, and interference environment detection. However, these methods still face several limitations. Current disease and pest detection models often rely on adaptive feature fusion, attention mechanisms, and feature guidance to improve performance. Despite their benefits, adaptive feature fusion can introduce biases in feature selection, leading to the omission of important information. Meanwhile, attention mechanisms tend to require substantial computational resources, which can adversely affect real-time performance. Feature guidance methods are usually tailored to specific features and thus lack the adaptability needed for detecting various types of plant diseases. For small object detection, techniques typically focus on enhancing image contrast and combining low-level and high-level

features to boost accuracy. However, in complex backgrounds, excessive contrast enhancement may cause confusion between the target and the background. Moreover, the process of aggregating low-level and high-level information is susceptible to interference from the background, limiting the model's ability to accurately detect small objects. Additionally, existing algorithms designed to operate in adverse weather conditions generally suffer from low detection accuracy, which restricts their practical application. Therefore, there is an urgent need for a highly accurate and robust object detection algorithm that can effectively reduce the impact of complex interference in challenging environments.

III. Dataset Acquisition and Interference Mechanism Injection

A. Data Set Acquisition

The study area is situated in Longyou County, Quzhou City, Zhejiang Province, China, with geographical coordinates of 119.17°E and 29.02°N. The study area comprises four towns in Longyou County: Hengshan, Shifo, Zhaxi, and Xiaohanhai, covering a total area of approximately 8 square kilometers. The main vegetation coverage types, according to China's secondary forest resource survey, are masson pine forests, broad-leaved evergreen forests, and coniferous and broad-leaved mixed forests. Masson pine forests account for a major proportion, ranging from 70% to 100% of the trees. The density of discolored pine trees is highest in Zhaxi Town, as shown in Fig. 1, which displays the study area and sampling plots. A total of seven plots were collected, including Plot 1 and Plot 2 in Tianchi Village, Hengshan Town, and Yanglong Reservoir, Plot 3 in Fengtang Village, Shifo Town, and Plot 4 and Plot 5 in Xihai Village, Hengshan Town, and Baiheqiao Village. Plot 6 is situated in Zhesi Village, Zhaxi Town, and Plot 7 is situated in Guiguangyan Village, Xiaonanhai Town.

TABLE I
DATASET SEGMENTATION

Dataset	Sampling plots	Dead infected pine trees
Training set	1,2,3,6	3,488
Validation set	7	501
Test set	4,5	997
Total	1,2,3,4,5,6,7	4,986

During data collection, we meticulously schedule our activities to align with the optimal time window for monitoring pests and diseases. Most pine trees infected with PWD exhibit typical reddish brown symptoms [26]. In contrast, the broad-leaved trees, which constitute a small proportion of the forest, have not yet changed color. Therefore, they will not be mistaken for the discolored pine trees. To ensure the accuracy of labeling and the credibility of testing, we have invited on-site investigation experts to provide us with guidance during the labeling process. We labeled the pine line pests in the images for this dataset using a labeling tool called LabelImg. Firstly, we selected the corresponding infected objects from the dataset images using rectangular boxes. Then we selected the pest marker category from the box and saved the marked
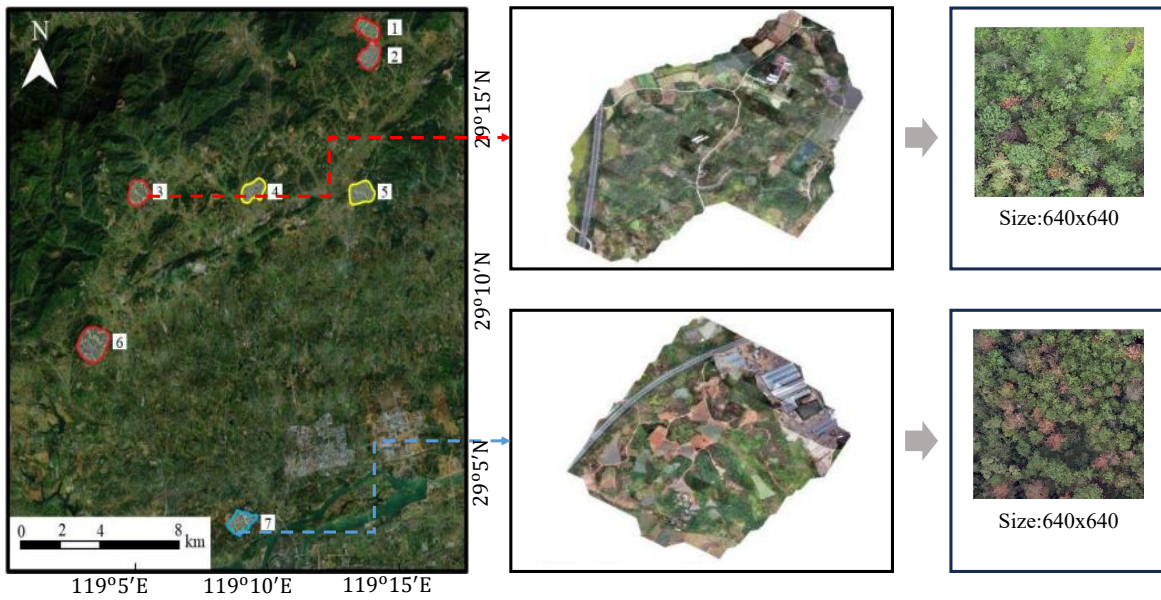
Fig. 1. Study area and sampling site.

label as a txt format label file. Finally, the labeled dataset was split into training, validation, and test sets, as detailed in Table I.

B. Interference With the Injection Mechanism

When remote sensing images are affected by atmospheric scattering, cloud and fog occlusion, and changes in lighting conditions, they can introduce random noise and result in image blurring. We hypothesize that the size and density of fog droplets follow a uniform distribution, allowing us to approximate the noise produced by fog interference as Gaussian noise. Gaussian noise, a well-known type of random noise with a normal distribution [27], is prevalent in image processing applications. The formula for generating Gaussian noise is as follows:

$$N\left(\mu, \sigma^2\right) = \mu + \sigma \times Z \tag{1}$$

where $N\left(\mu, \sigma^2\right)$ represents a Gaussian. The symbols μ and σ^2 represent the mean and variance of the Gaussian distribution, respectively. Notably, the variance σ^2, dictates the intensity of the noise introduced, with larger values indicating stronger noise corruption. Z is the random number following the standard normal distribution. The interference sample can be obtained by superposing the generated Gaussian noise into the original PWD dataset sample. The interference sample generation process under Gaussian noise interference is shown in Fig. 2.

We treat different degrees of fog interference as Gaussian noise interference of different intensities. The variance of the equivalent Gaussian noise increases as fog intensity rises in the image. The intensity of Gaussian noise $\sigma \in \{0.02, 0.03, 0.05, 0.07\}$ corresponds to the interference of light fog, medium fog, heavy fog, and dense fog on the remote sensing images.
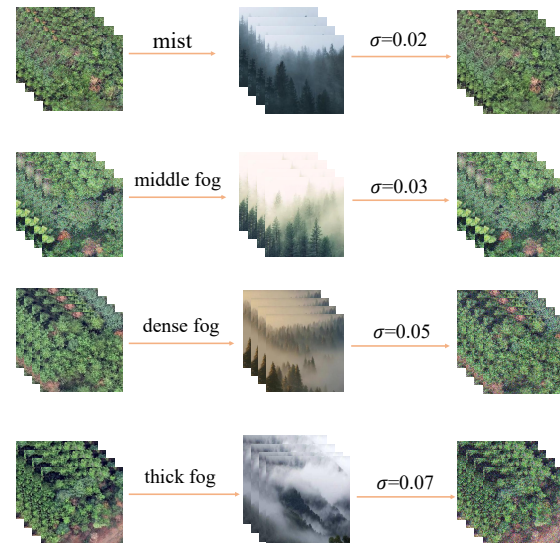

Fig. 2. Interference sample generation process.

Poisson noise is caused by the randomness of photons during the imaging process, and applying Poisson noise can reflect the random distribution of light under different illumination conditions. By adjusting the parameters of Poisson noise, it is possible to simulate infected pine forests under various lighting conditions. The formula for generating Poisson noise can be expressed as follows:

$$P(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{2}$$

where $P(k; \lambda)$ represents the probability of an event occurring k times. λ denotes the mean of the Poisson distribution, $e^{-\lambda}$ represents the monotonically decreasing exponential function, and k indicates the actual number of occurrences of the event.
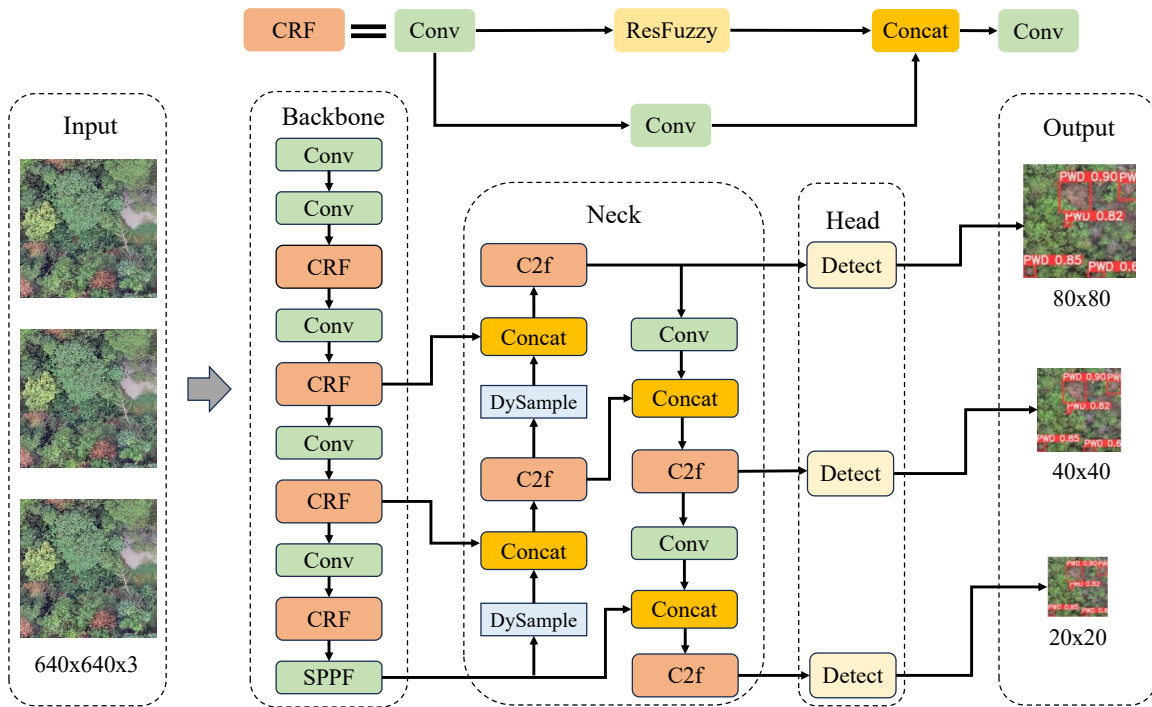
Fig. 3. The structure of YOLOv8-RD network.

After generating Poisson noise, it is then superimposed onto the original PWD images to create a dataset with different lighting conditions. We subsequently employed a sample dataset of PWD with Gaussian noise of varying interference intensities to train and test the anti-interference performance of the object detector.

IV. YOLOV8-RD

In this study, we adopted the YOLOv8n version as the basis of our detection model due to its well-balanced trade-off between detection speed and accuracy. To achieve robust detection of PWD in remote sensing images under various interference conditions, this paper proposes an improved YOLOv8n model, named YOLOv8-RD, as shown in Fig. 3. We combine the advantages of residual learning with FDNN, designed a ResFuzzy module to effectively learn image residuals and suppress noise over-expression. Meanwhile, the ResFuzzy module integrates a DPM feature enhancement module to strengthen global information extraction, consequently improving the model's ability to detect small objects. Additionally, the conventional nearest neighbor interpolation is replaced with a DySample module in this study. By dynamically adjusting the sampling points, DySample module enhances the sensitivity and adaptability of the sampling process to input feature variations, thereby improving the robustness of the object detection model under noise.

A. Residual Fuzzy ResFuzzy

During remote sensing image capture, UAV can be influenced by atmospheric scattering, cloud and fog occlusion, and fluctuations in lighting conditions. These factors can lead to image blurring and markedly impair the detection capability of object detection models.

To raise the anti-interference ability of object detection models in complex interference environments, this study proposes ResFuzzy module, as shown in Fig. 4. This module can effectively suppress the over-expression of noise and improve the robustness of the object detection model.

The ResFuzzy module consists of two sub-residual blocks, a DPM module [28], and a Fuzzy layer [29]. In Fig. 4, the input features x are transmitted through the first sub-residual block to obtain residual feature values y_1. Through the DPM module, the low-frequency detail information transmitted in residual learning is effectively enhanced, and high-frequency noise is filtered out, resulting in the output y_2. Next, the output y_2 is input into the second sub-residual block to obtain the residual feature value y_3, which is then input into the fuzzy layer. By smoothing the feature map values using the Gaussian function, the blurred output y_4 is obtained. Finally, the obtained fuzzy output is input into the C3 module and fused with the input noisy image through residual connection to obtain a high-resolution image while filtering out image noise. We renamed the improved C3 module CRF, as shown in Fig. 5.

During the convolution process, the noise will propagate downwards with the convolution, resulting in a decreased image resolution and an increased difficulty in detection [30]. Therefore, this study introduces residual learning to directly transfer shallow information to deep layers to filter out image noise [31, 32]. The two residual blocks of the ResFuzzy module are composed of a 1×1 convolutional layer, a linear rectification activation layer, and a 3×3 convolutional layer. This design helps extract image features through skip layer connections and suppress over-expression of noise [33, 34].
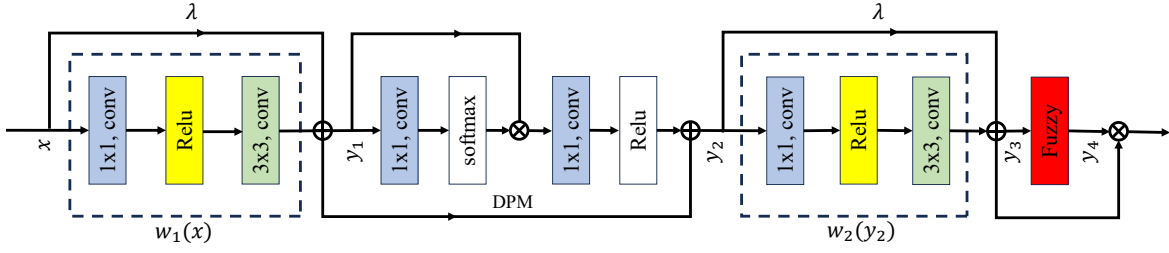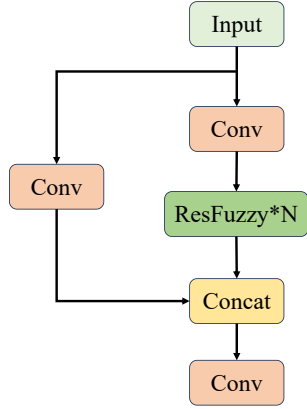
Fig. 4. The structure of ResFuzzy module.

Fig. 5. The structure of CRF module.

The output of two sub-residual blocks can be represented as:

$$y_1 = w_1(x) + \lambda x \qquad (3)$$

$$y_3 = w_2(y_2) + \lambda y_2 \qquad (4)$$

where $w_1(x)$ represents the residual mapping learned by the first sub-residual block, λ represents the degree of preservation of the original input, and $w_2(y_2)$ represents the residual mapping learned by the second sub residual block.

The DPM module can capture long-range dependencies, obtain contextual information, and enhance the low-frequency detail information transmitted in residual learning [35, 36]. Its model is defined as:

$$\mathrm{DPM}(y_1) = y_1 + \gamma\left(\mathrm{F}(\hat{y_1})\right) \qquad (5)$$

where y_1 represents the input features. $\hat{y_1} = \sigma(F(y_1)) \times y_1$, F is the convolutional layer with kernel 1×1, γ is the LeakyReLU activation function, and σ is the Softmax function.

After the input features y_1 enter the DPM module, they go through a 1×1 convolutional layer and a $Softmax$ activation layer to obtain a normalized spatial attention mask. This mask is then multiplied with the input features to obtain global context information. The obtained global context information is then input into a sub-network consisting of a 1×1 convolutional layer and a $ReLU$ activation function to obtain channel attention features. The input features y_1 are then added to the obtained channel attention features via a residual connection to get the enhanced detailed features.

FDNN can significantly suppress excessive noise expression through the processes of fuzzification and defuzzification, thereby providing neural networks with stronger robustness and generalization ability [37]. To further bolster the robustness of object detection networks in complex interference environments, we incorporate a fuzzy layer in our improved residual network.

In this study, we introduce trainable fuzziness parameters d and standard deviation parameters σ into the fuzzy layer to compute the difference between each channel and the fuzziness parameter. The generated differences are then fuzzified using a Gaussian function to smooth out noise in the input data, ensuring that variations in input features within a certain range do not significantly affect the output results [38]. The fuzzification process is described in Equation 6.

$$fuzzy\_out = \exp\left(-\sum_{i=1}^{c} \frac{(y_i - d_i)^2}{\sigma_i^2}\right) \qquad (6)$$

where $fuzzy\_out$ represents the fuzzy output. exp represents the exponential function. c denotes the number of channels. y_i is the output of the $i-th$ channel in the improved residual block. d_i is the fuzziness parameter for the $i-th$ channel, and σ_i is the standard deviation parameter for the $i-th$ channel. Upon processing the feature map through the fuzzy layer, it is subjected to element-wise multiplication with the output of the refined residual block. This ensures that each channel undergoes noise suppression and smoothing, thus enhancing the robustness of the model against complex interference environments.

B. Dynamic Upsampling Operator DySample

Feature upsampling plays a crucial role in progressively restoring feature resolution in dense prediction tasks [39]. Nearest neighbor interpolation is often used for feature upsampling in YOLOv8n. However, this upsampling method copies the values of the nearest pixels to restore the feature resolution [40]. If the image itself contains noise, the noise will be amplified after upsampling. To more effectively restore image features under noise interference, we replace the upsampling method of nearest neighbor interpolation with the DySample module [41].

The DySample module is an upsampling method based on point sampling, which differs from traditional fixed-rule and kernel-based dynamic upsampling techniques. By learning

the offset of input feature points, the DySample module dynamically determines the upsampling position, thereby realizing content-aware upsampling behavior. The DySample module mitigates noise amplification issues inherent in nearest neighbor interpolation and effectively reduces the impact of image noise during the upsampling process, thereby enhancing the quality of the upsampled feature map [42].

(a) Sampling based dynamic upsampling
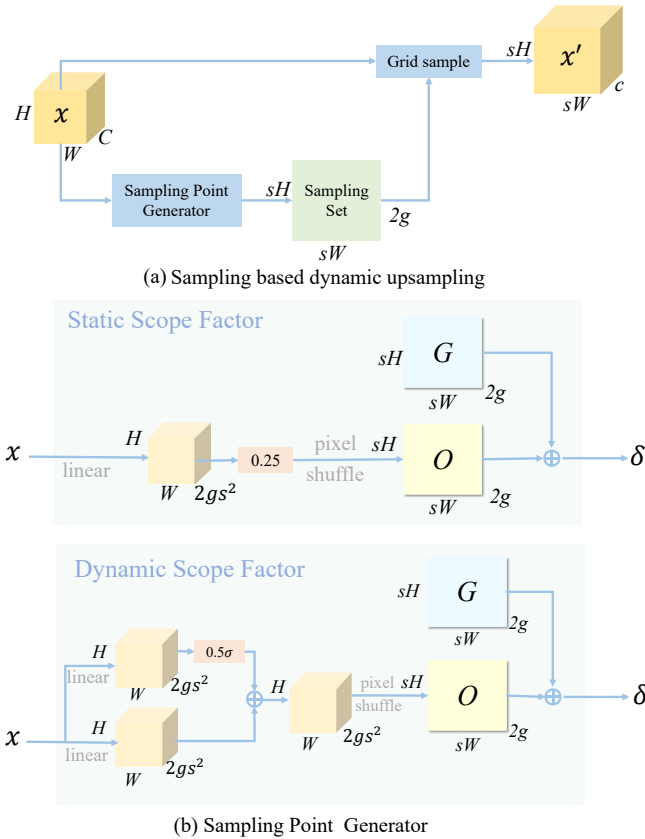
(b) Sampling Point Generator

Fig. 6. The structure of DySample module.

Fig. 6(a) shows the process of dynamic upsampling using DySample. The feature map x with dimensions $H \times W \times C$ is input into the DySample module. x is processed by the sampling point generator to learn feature offsets. These offsets are then used to create a sampling set with dimensions $sH \times sW \times 2g$. The grid sampling function then resamples the input feature x using these offsets, resulting in an upsampled feature map x' with dimensions $sH \times sW \times C$. To effectively learn the offsets for different feature points, we designed two sampling point generators, as shown in Fig. 6(b). Here, s represents the scale factor for upsampling, g represents the number of groups respectively, and x represents the input feature. $linear$ denotes the linear layer, and σ denotes the sigmoid activation function. O represents the generated offset, G represents the original sampling grid, δ represents the generated sampling set, and s^2 represents the number of repetitions of the offset in each dimension.

Fig. 6(b) illustrates two methods for generating sampling points, static scaling factor and dynamic scaling factor. The static scaling factor uses a fixed factor of 0.25 to limit the range of offsets, and combines a linear layer and pixel shuffle

to generate the offset O. The offset O is then added to the original sampling grid G to obtain the sampling set δ. The static sampling point generator can limit the local spatial range that each upsampled point can traverse, ensuring that the sampling positions do not overlap excessively, avoiding blurred boundaries and error propagation in the output feature map.

The dynamic scaling factor is generated dynamically based on the input features, allowing it to offer varying offset ranges for each sampling point [43]. The input feature map x enters the dynamic sampling point generator and is processed through two branches. The first branch is processed by a linear layer $linear_1$, which maps each channel of the feature map to a new feature space, thereby converting each feature point into an offset. Then, the Sigmoid function is used to map the offset to the interval [0,1], and the obtained value is multiplied by 0.5 to derive a dynamic scaling factor centered at 0.25. The second branch is processed by a linear layer $linear_2$ to directly generate the offset O. The generated dynamic scaling factor is used to adjust the offset produced by $linear_2$, thereby enabling dynamic adjustment of the offset. The offset adjustment process can be represented by Equation 7.

$$O = 0.5\text{Sigmoid}\left(\text{linear}_1(x)\right) \cdot \text{linear}_2(x) \qquad (7)$$

The dynamic range factor allows for adaptive adjustment of the step size in the sampling point movement based on content variations of the input features. This enhances the sensitivity and adaptability of the sampling process to changing features, paving the way for a more robust object detection model against noise.

V. Experimental Results and Analysis

A. Experimental Setup

TABLE II
TRAINING PARAMETERS

Training Parameters	Details
Epochs	130
Image size (Pixel)	640×640
Batch size	16
Workers (Number of Threads)	16
d (Fuzziness Parameters)	[0,1]
σ (Standard Deviation)	1
Initial learning rate	0.01
Final OneCycleLR learning rate	0.1
Optimization algorithm and parameters	Adam(0.937)
Weight decay	0.0005

In the experiments, we initialized the parameters of the YOLOv8n network before training on the PWD dataset. The hyperparameters of the experimental model are detailed in Table II. Additionally, Table III outlines the software and hardware configurations used throughout the experiments. Both the original YOLOv8n and the YOLOv8-RD network models proposed in this paper were trained based on the pre-trained model (YOLOv8n.pt). During the training process, the convergence

TABLE III
EXPERIMENTAL ENVIRONMENT

Experimental Environment	Details
Programming language	Python 3.8
Operating system	Windows 10
Deep learning framework	Pytorch 1.11.0
CPU	22 vCPU AMD EPYC 7T83 64-Core
GPU	NVIDIA GeForce RTX 3090

point of each model was recorded at its peak performance, and its efficacy was assessed using a dataset with varying levels of interference strength and different interference ratios. In this study, Precision (P), Recall (R), and mean Average Precision (mAP) are used as evaluation metrics for the model. Higher P indicates a lower false positive rate, which reflects the model's ability to mistakenly detect background interference as diseased objects. High R means the model can detect as many objects as possible in a noisy environment. mAP provides a more comprehensive reflection of the robustness of the object detection model in noisy environments. A higher mAP indicates a superior anti-interference capability of the model. The P, R and mAP calculation formula is as follows:

$$P = \frac{TP}{TP + FP} \tag{8}$$

$$R = \frac{TP}{TP + FN} \tag{9}$$

$$AP = \int_0^1 PdR \tag{10}$$

$$mAP = \frac{1}{n} \sum_{i=0}^{n} AP_i \tag{11}$$

where P represents the detection accuracy, which is the proportion of correctly detection PWD relative to the total number of objects the model detected as PWD. R represents the recall rate, which is the proportion of correctly detected PWD compared to the total number of PWD present in the dataset. TP denotes the number of correctly detected PWD. FP denotes the number of falsely detected PWD. FN denotes the number of falsely detected non-PWD. AP denotes the average highest precision for different object categories. mAP denotes the mAP across all object categories. n denotes the number of detected objects, and i denotes the object currently being detected.

B. Ablation Experiments

To assess the impact of the two additional modules in the YOLOv8-RD model on the original YOLOv8n model, we conducted four groups of ablation experiments. The following modifications are made on the basis of YOLOv8n. We integrated the ResFuzzy module into the backbone network and implemented the DySample module in the neck network. We used training samples with a Gaussian noise intensity of 0.02 for adversarial training and applied the resulting model weights in ablation experiments. The ablation experiments

were conducted using a test set with 100% noisy samples, and the experimental results are presented in Table IV. Additionally, the changes in the model parameters are shown in Table V. Where, Parameters refers to the number of parameters in the model during training, measured in Millions (M). Increment indicates the increase in the number of parameters after adding different modules to the model.

TABLE IV
RESULT OF ABLATION EXPERIMENTS

ResFuzzy	DySample	P(%)	R(%)	mAP(%)
-	-	84.3	78.9	87.2
√	-	86.7	83.4	90.4
-	√	83.6	82.1	88.2
√	√	86.5	85.9	91.1

TABLE V
MODEL PARAMETERS

Model	Parameters (M)	Increment (M)
YOLOv8	3.011	0
YOLOv8+ResFuzzy	3.065	0.054
YOLOv8+DySample	3.023	0.012
YOLOv8+ResFuzzy+DySample	3.077	0.066

As shown in Table IV, before adding the anti-interference module, the average precision of the YOLOv8n model was only 87.2%. With the introduction of the ResFuzzy module, the accuracy, recall, and average precision of the model increased by 2.4%, 4.5%, and 3.2%, respectively. This indicates that the ResFuzzy module can effectively filter out noise in images through multi-level residual blocks and fuzzy layers under Gaussian noise interference, thereby improving the model's detection accuracy in noisy environments. In noisy environments, recall may decrease as the model might overlook some objects whose features are obscured by interference. Nonetheless, with the incorporation of the DySample module, while accuracy undergoes a minor decline, both recall and average precision exhibit a subtle enhancement. This indicates that the DySample module can adjust the sampling point step size according to the changes in the input image noise, effectively restoring the object features of the input feature map and improving the quality of the upsampled feature map. With the addition of both modules, the number of parameters of the model only increases slightly, nevertheless, the accuracy, recall, and average precision of the model are significantly improved. This fully demonstrates that the two modules we added can effectively resist noise and achieve high-precision detection in noisy environments.

C. Comparative Analysis of Different Interference Intensities

When a UAV captures images of a vast pine forest along a designated path, noticeable cloud and fog occlusion may occur in certain regions, leading to the blurring of disease features in some datasets. Therefore, we created test sets with interference sample proportions of 10%, 30%, 50%, 80%, and 100% to

simulate the situation where different numbers of samples in the dataset are affected by fog. At the same time, we used a Gaussian noise intensity of $\sigma \in \{0.02, 0.03, 0.05, 0.07\}$ to simulate the interference of different fog intensities.

We conducted preliminary comparisons between our proposed YOLOv8-RD model and traditional models such as YOLOv5s, YOLOv6, YOLOv7-tiny, and YOLOv8n. Additionally, to investigate the superiority of our model, we further compared YOLOv8-RD with four most representative models: YOLOv9-t, YOLOv10n, RTDETR-r18, and SC-RTDETR. Among them, the SC-RTDETR model is the latest object detection model proposed by Feng H et al. [44], which features strong interference robustness. Tables VI-IX present the mAP variations of different models on the training and test sets under interference intensities of 0.02, 0.03, 0.05, and 0.07, respectively. Where Tr-Ac represents the accuracy of the training set, and Te-Ac represents the accuracy of the interference-free test set. The entire experiment adopts an adversarial training method to train the weights of different models under different interference intensities, and then uses the trained weights to test the anti-interference ability of the test set without interference and the test set with samples of different interference proportions.

TABLE VI
INTERFERENCE INTENSITY σ=0.02

Model	Tr-Ac	Te-Ac	10%	30%	50%	80%	100%
YOLOv5s	92.0	89.1	88.7	88.3	87.4	86.9	86.7
YOLOv6	85.3	84.1	83.5	82.0	81.9	81.5	80.9
YOLOv7-tiny	89.4	88.7	87.9	86.9	86.2	85.6	85.2
YOLOv8n	90.5	89.7	89.1	88.8	88.5	87.6	87.2
YOLOv9-t	90.7	90.5	90.2	89.7	88.4	87.9	87.4
YOLOv10n	86.2	83.2	82.5	83.6	83.8	85.3	85.2
RTDETR-r18	91.8	88.6	87.9	86.9	86.8	87.1	87.3
SC-RTDETR	92.3	91.5	91.3	90.3	89.8	89.5	87.9
YOLOv8-RD	92.5	92.2	92.0	91.9	91.7	91.3	91.1

TABLE VII
INTERFERENCE INTENSITY σ=0.03

Model	Tr-Ac	Te-Ac	10%	30%	50%	80%	100%
YOLOv5s	89.5	88.5	88.3	87.8	86.9	85.8	85.4
YOLOv6	83.4	82.9	82.5	81.7	80.6	79.5	77.6
YOLOv7-tiny	88.9	88.3	87.7	86.4	85.1	84.7	84.0
YOLOv8n	89.8	88.7	88.2	87.5	87.3	86.5	85.2
YOLOv9-t	91.6	91.0	90.1	89.4	89.0	87.5	86.9
YOLOv10n	85.8	82.4	82.8	82.4	84.7	85.2	84.9
RTDETR-r18	87.7	82.7	81.5	81.6	83.6	84.6	86.3
SC-RTDETR	91.4	85.5	85.8	86.1	87.9	88.8	87.1
YOLOv8-RD	92.0	91.5	91.3	91.1	90.8	90.4	90.2

Without interference, the average precision of the YOLOv8-RD model on the test sets with various interference intensities has consistently remained above 90%, with a maximum precision of 92.2%. This signifies that the YOLOv8-RD model not only effectively detects object features amidst interference but also maintains exceptional detection accuracy in unobstructed environments, thereby providing a robust solution for the efficient detection of PWD.

Under interference conditions, all models experienced varying degrees of accuracy decline, especially when the interference ratio in the test set reached 100% and the interference

TABLE VIII
INTERFERENCE INTENSITY σ=0.05

Model	Tr-Ac	Te-Ac	10%	30%	50%	80%	100%
YOLOv5s	87.6	86.9	85.2	84.3	83.1	82.2	80.2
YOLOv6	82.8	81.7	81.2	80.6	77.8	75.1	74.3
YOLOv7-tiny	87.4	86.3	84.5	83.2	81.5	80.8	79.6
YOLOv8n	89.2	87.5	85.6	84.9	83.5	81.7	80.4
YOLOv9-t	92.3	91.5	90.3	87.9	85.3	84.1	83.3
YOLOv10n	84.8	84.5	83.5	83.5	81.7	81.4	79.1
RTDETR-r18	89.7	85.9	84.6	83.8	83.6	83.4	79.5
SC-RTDETR	91.8	89.4	88.4	87.2	86.6	83.7	83.3
YOLOv8-RD	91.8	91.5	91.0	90.2	90.1	89.1	87.2

TABLE IX
INTERFERENCE INTENSITY σ=0.07

Model	Tr-Ac	Te-Ac	10%	30%	50%	80%	100%
YOLOv5s	86.2	84.8	84.2	82.6	81.1	78.3	77.2
YOLOv6	80.6	78.7	77.5	76.9	75.3	73.4	72.3
YOLOv7-tiny	86.5	84.8	84.1	82.0	80.8	76.6	75.3
YOLOv8n	87.2	85.2	84.8	83.5	80.6	78.9	77.8
YOLOv9-t	91.5	90.8	87.2	86.5	84.6	82.9	81.1
YOLOv10n	85.4	82.6	82.6	82.5	82.8	79.6	79.7
RTDETR-r18	88.8	82.0	81.3	79.6	78.4	80.4	78.7
SC-RTDETR	90.5	88.6	87.5	85.7	86.2	83.3	83.0
YOLOv8-RD	90.8	90.7	90.4	90.2	88.2	87.3	86.0

intensity reached 0.07, where the accuracy of each model dropped most considerably. The YOLOv6 model exhibited the most pronounced average accuracy fluctuations, whereas the YOLOv8n and YOLOv10n models demonstrated commendable robustness under lower interference intensities. The YOLOv9-t model exhibited high accuracy in both the training set and the interference-free test set, but its detection accuracy dropped rapidly when exposed to noise. Both the SC-RTDETR and YOLOv8-RD models showed strong robustness under various interference conditions. However, the YOLOv8-RD model performed significantly better than the SC-RTDETR model under various levels of interference, with a maximum accuracy improvement of 5.5%. These results indicate that the YOLOv8-RD model exhibits greater robustness and higher detection precision in complex interference environments. Such phenomenon occurs when Gaussian noise obscures the details and texture information in images, leading to an increase in image blurriness, which in turn makes it challenging to detect and extract features from the image, thereby affecting detection accuracy. In contrast, the proposed YOLOv8-RD model shows significant advantages in this perspective. On the one hand, the ResFuzzy module filters and smooths the image when encountering noisy images, thereby mitigating the impact of noise on the model. On the other hand, the DySample module adjusts the sampling points in real-time according to changes in input feature noise intensity, enhancing the model's robustness in complex interfering environments.

D. Comparative Analysis of Different Interference Methods

When using UAV to monitor pine forests for pests and diseases, two main environmental disturbances emerges: cloud cover and variations in lighting. Cloud cover can cause image blurriness and reduced contrast, while changes in lighting can lead to shadows and non-uniform brightness in images [45].
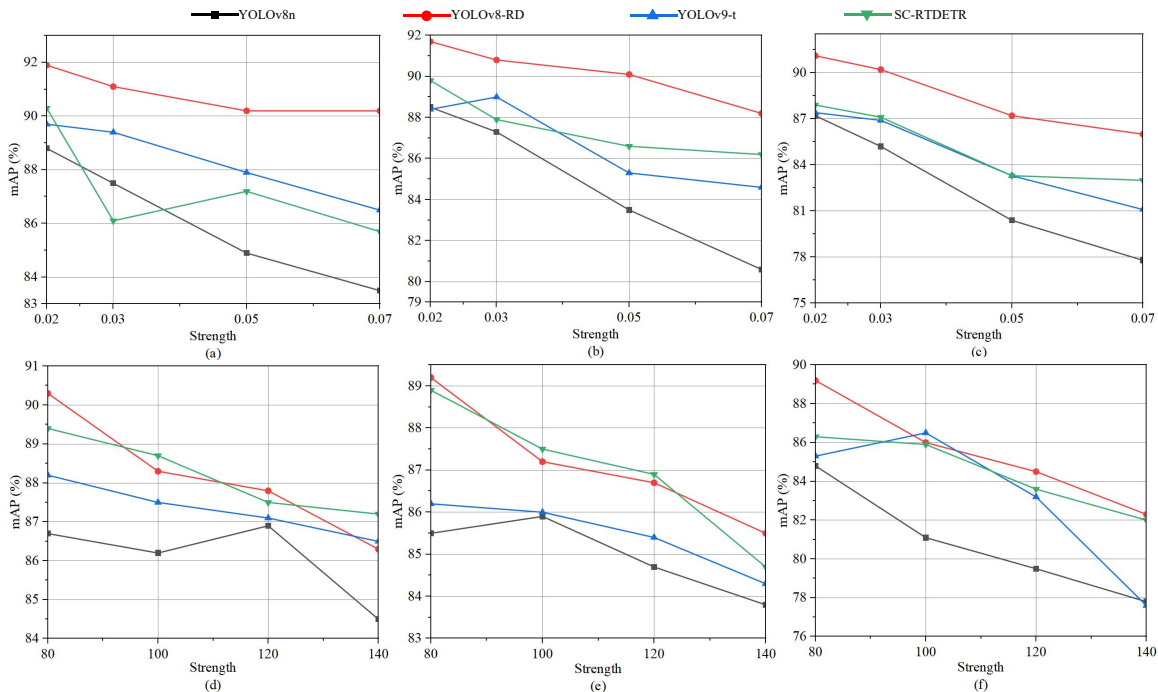
Fig. 7. Comparative experiment of different interference methods.

These disturbances significantly impact the feature extraction process of object detection models, making it difficult for the algorithm to accurately locate diseased areas, resulting in false positives and false negatives. To address these issues, we employ Gaussian noise to simulate cloud cover and Poisson noise to simulate variations in lighting.

In the experiments, we selected a Gaussian noise intensity of $\sigma \in \{0.02, 0.03, 0.05, 0.07\}$ and a Poisson noise intensity of $\lambda \in \{80, 100, 120, 140\}$. To verify the effectiveness of our proposed YOLOv8-RD for both types of interference, we selected three of the most representative models for comparison, YOLOv8n, YOLOv9-t, and SC-RTDETR. At the same time, the entire experiment adopts an adversarial training method, with an adversarial sample ratio of 30% in the training set. We use the trained model weights to test the anti-interference ability of the test set with different interference proportions. The test results are shown in Fig. 7. Specifically, Figs. 7(a), 7(b), and 7(c) display the results of different models' anti-interference tests on the test sets with Gaussian noise interference proportions of 30%, 50%, and 100%, respectively. Figs. 7(d), 7(e), and 7(f) show the results of different models' anti-interference tests on the test sets with Poisson noise interference proportions of 30%, 50%, and 100%, respectively.

As shown in Figs. 7(a), 7(b), and 7(c), when subjected to Poisson noise, the proposed YOLOv8-RD model exhibits significantly better anti-interference performance than the traditional YOLOv8n model in tests with various interference ratios. Moreover, in most cases, the YOLOv8-RD model outperforms the state-of-the-art YOLOv9-t model in terms of anti-interference performance. Finally, the YOLOv8-RD model is compared to the SC-RTDETR model, which also has an anti-interference module. The proposed YOLOv8-RD model has a slightly lower mAP than the SC-RTDETR model in the test sets with interference ratios of 30% and 50%. However, when the interference sample ratio is 100%, the YOLOv8-RD model exhibits higher mAP under different interference intensities. This indicates that the proposed YOLOv8-RD model can effectively resist interference caused by changes in lighting, reducing false positives and false negatives caused by factors such as exposure and reflection.

By observing Figs. 7(d), 7(e), and 7(f), it can be deduced that the proposed YOLOv8-RD model has significantly higher mAP in detecting objects under Gaussian noise interference compared to the other three models. This indicates that YOLOv8-RD can effectively filter out Gaussian noise and accurately detect diseased objects. In contrast, the performance of the other three models decreases significantly when subjected to Gaussian noise interference of different intensities and proportions.

Two sets of comparative experiments both point out that the proposed YOLOv8-RD model can maintain high detection accuracy and strong robustness in the case of cloud and fog occlusion and changes in lighting. On the one hand, this is due to the fact that the ResFuzzy module can use multi-level residual blocks to filter image noise and combine a fuzzy layer to suppress excessive noise expression. On the other hand, the DySample module effectively improves the quality of the upsampled feature map by dynamically adjusting the sampling points. The combination of the two modules significantly improves the anti-interference performance of the YOLOv8n model.

E. Analysis of Grad-CAM

To better demonstrate the advantages of the YOLOv8-RD object detection model on the interfered dataset, the Gradient-

weighted Class Activation Mapping (Grad-CAM) technique is used to visualize the regions of interest of the model in the image and enhance the interpretability and reliability of the model [46]. Grad-CAM generates a coarse heat map by multiplying the gradients of the output class with the outputs of a specific convolutional layer and then averaging the results to show the areas of the image the model is focusing on [47]. By analyzing the generated heat maps, we can detect potential shortcomings where the model might focus on irrelevant image areas or miss crucial features for accurate object detection. Its formula can be simply expressed as:

$$L_{\text{Grad-CAM}}^{c} = \text{ReLU}\left(\sum_{i}\sum_{j}\frac{\partial Y^{c}}{\partial A_{ij}}A_{ij}\right) \qquad (12)$$

where Y^{c} represents the output score of the model for a specific class, $\frac{\partial Y^{c}}{\partial A_{ij}}$ represents the gradient of that class score with respect to the feature map A, and A_{ij} represents the element at the i-th row and j-th column of feature map A. By summing up the weights of each channel and applying the ReLU activation function, a non-linear mapping is obtained.

We selected the 15-th layer of the model for backpropagation and generated heat maps by applying Grad-CAM to two images with Gaussian noise intensity of 0.02. The heat maps allow us to observe the changes in the regions of interest of the YOLOv8n and YOLOv8-RD models in noisy environments. The comparison are shown in Fig. 8.
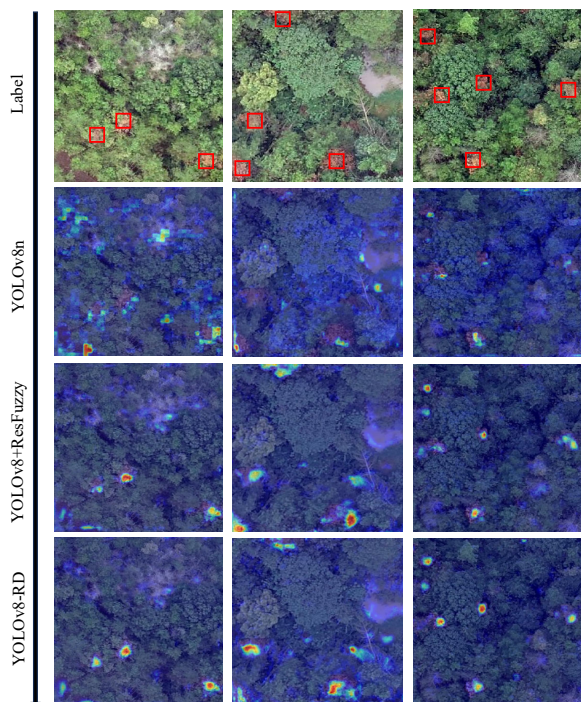

Fig. 8. Gradient-weighted class activation mapping comparison.

In the generated heat maps, deeper red shadows indicate regions of high attention from the model. Yellow regions represent diseased object areas with lower attention, and blue regions indicate areas with redundant and interfering information. In the heat map generated using the YOLOv8n

model weights, there are obvious blue interference regions, and the key red object regions are fewer. This suggest that the traditional YOLOv8n model have difficulty filtering out noise in noisy environments, thereby impeding its ability to effectively differentiate between object and background features. However, with the addition of the ResFuzzy module, the blue regions in the heat map are significantly reduced, and the color of the key red object regions is significantly deepened. This demonstrates that the ResFuzzy module can markedly improve the resolution of object detail features while effectively mitigating noise. After incorporating the DySample module, the blue regions in the heat map are further reduced, and the red and yellow regions are further enhanced. This indicates that the DySample module has a certain inhibitory effect on noisy environments, making the model more sensitive to changes in noise in the feature map.

F. Anti Interference Test Results

To straightforwardly demonstrate the effectiveness and advantages of the YOLOv8-RD model in interfering environments, we conducted four adversarial training experiments on datasets with two different levels of Gaussian and Poisson noise. The trained model weights were then used to evaluate the detection performance of both YOLOv8n and YOLOv8-RD on single images. In our experiments, we selected test samples with two different noise levels from both Gaussian and Poisson noise datasets to assess the models' robustness to noise. The detection results are shown in Fig. 9.
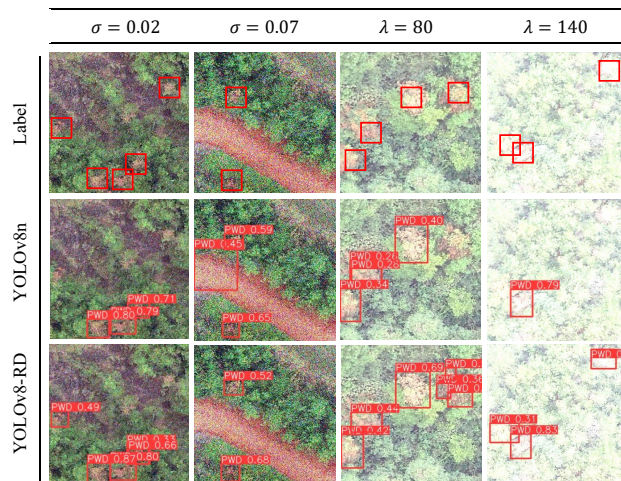

Fig. 9. Anti interference test results.

As shown in the figure above, under the interference of Gaussian noise, the original YOLOv8n model encountered issues with missed detection and false positives. In environments with substantial noise, the features of small objects may become obscured, thereby impeding the YOLOv8n model's capacity to discern these minute details with precision. Additionally, similar ground features were mistakenly detected as disease objects by the YOLOv8n model. In contrast, the proposed YOLOv8-RD model was able to accurately detect nearly all object features under the simulated cloud and fog interference from Gaussian noise.

When exposed to the interference of illumination changes simulated by Poisson noise, the detailed features and edge contours of the objects in the image became blurred. In such environments, the YOLOv8n model struggled to accurately detect disease features, often resulting in missed detection. However, our proposed YOLOv8-RD model effectively resisted the interference from illumination changes and cloud cover. By leveraging the ResFuzzy module, image noise induced by cloud cover was effectively mitigated, while the DySample module facilitated the restoration of disease features that were compromised due to variations in illumination. The incorporation of these two anti-interference modules substantially bolstered the robustness of the YOLOv8n model, markedly decreasing both missed detection and false positive rates in challenging, noisy environments.

VI. Acknowledgment

VII. Conclusion

This study simulates cloud occlusion and lighting variations using Gaussian and Poisson noise. By designing the ResFuzzy module to suppress the interference of image noise and background features, we significantly improved the model's robustness in various noisy environments. In addition, we integrated the DySample module into the Neck part, effectively enhancing the adaptability of the upsampling process to changes in input features, further strengthening the model's resistance to interference.

Through extensive experiments, our proposed YOLOv8-RD model achieves a maximum detection accuracy of 92.2% in noise-free conditions. Under Gaussian and Poisson noise interference, the YOLOv8-RD model significantly outperforms four state-of-the-art models. This illustrates that the YOLOv8-RD model adeptly counters noise interference, facilitating precise detection of PWD in intricate and obstructive environments when deployed via UAV systems.

References

[1] A. K. Rana, S. Guleria, V. K. Gupta, and V. K. Thakur, "Cellulosic pine needles-based biorefinery for a circular bioeconomy," *Bioresource Technology*, vol. 367, p. 128255, 2023.

[2] L. Wang, J. Cai, T. Wang, J. Zhao, T. R. Gadekallu, and K. Fang, 'Pine Wilt Disease Detection based on UAV Remote Sensing with an Improved YOLO Model," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024. DOI: 10.1109/JS-TARS.2024.3478333.

[3] X. Lu, J. Huang, X. Li, G. Fang, and D. Liu, 'The interaction of environmental factors increases the risk of spatiotemporal transmission of pine wilt disease," *Ecological Indicators*, vol. 133, p. 108394, 2021.

[4] H. Li, L. Chen, Z. Yao, N. Li, L. Long, and X. Zhang, "Intelligent identification of pine wilt disease infected individual trees using UAV-based hyperspectral imagery," *Remote Sensing*, vol. 15, no. 13, p. 3295, 2023.

[5] Z. Yu, H. Huang, W. Chen, Y. Su, Y. Liu, and X. Wang, 'Yolo-facev2: A scale and occlusion aware face detector," *Pattern Recognition*, vol. 155, pp. 110714, 2024.

[6] X. Wang and J. Liu, "Vegetable disease detection using an improved YOLOv8 algorithm in the greenhouse plant environment," *Scientific Reports*, vol. 14, no. 1, p. 4261, 2024.

[7] Y. Tian, S. Wang, E. Li, G. Yang, Z. Liang, and M. Tan, "MD-YOLO: Multi-scale Dense YOLO for small target pest detection," *Computers and Electronics in Agriculture*, vol. 213, p. 108233, 2023.

[8] S. An, X. Huang, L. Cao, and L. Wang, 'A comprehensive survey on image dehazing for different atmospheric scattering models," *Multimedia Tools and Applications*, vol. 83, no. 14, pp. 40963–40993, 2024.

[9] K. Bhalla, D. Koundal, B. Sharma, Y.-C. Hu, and A. Zaguia, 'A fuzzy convolutional neural network for enhancing multi-focus image fusion," *Journal of Visual Communication and Image Representation*, vol. 84, pp. 103485, 2022.

[10] A. Saffari, M. Khishe, and S.-H. Zahiri, 'Fuzzy-ChOA: an improved chimp optimization algorithm for marine mammal classification using artificial neural network," *Analog Integrated Circuits and Signal Processing*, vol. 111, no. 3, pp. 403–417, 2022.

[11] D. Pandey, B. K. Pandey, and S. Wairya, "Hybrid deep neural network with adaptive galactic swarm optimization for text extraction from scene images," *Soft Computing*, vol. 25, no. 2, pp. 1563–1580, 2021.

[12] M. Sibiya and M. Sumbwanyambe, 'Automatic fuzzy logic-based maize common rust disease severity predictions with thresholding and deep learning," *Pathogens*, vol. 10, no. 2, pp. 131, 2021.

[13] X. Tian, X. Meng, Q. Wu, Y. Chen, and J. Pan, 'Identification of tomato leaf diseases based on a deep neuro-fuzzy network," *Journal of The Institution of Engineers (India): Series A*, vol. 103, no. 2, pp. 695–706, 2022.

[14] A. K. Koshariya, M. S. Ashraf, R. Nigam, T. S. Kumar, D. Kumar Awasthi, and G. Elavarasan, 'Implementation of Deep Convolution Neuro-Fuzzy Network to Plant Disease Detection, Risk Assessment, and Classification," in *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, pp. 575–581, 2023.

[15] B. Liu, X. Bai, X. Su, C. Song, Z. Yao, X. Wei, and H. Zhang, "DAC-PPYOLOE+: A lightweight real-time detection model for early apple leaf pests and diseases under complex background," in *Proceedings of the 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 174–182, 2023.

[16] J. Qi, X. Liu, K. Liu, F. Xu, H. Guo, X. Tian, M. Li, Z. Bao, and Y. Li, "An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease," *Computers and Electronics in Agriculture*, vol. 194, p. 106780, 2022.

[17] F. Deng, W. Mao, Z. Zeng, H. Zeng, and B. Wei, "Multiple diseases and pests detection based on federated learning and improved faster R-CNN," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022.

[18] S. Y. Irianto and E. Findley, 'Fuzzy Deep Learning Recurrent Neural Network Algorithm to Detect Corn Leaf Disease," *International Journal of Computational Intelligence and Applications*, p. 2450017, 2024.

[19] B. Chang, Y. Wang, X. Zhao, G. Li, and P. Yuan, 'A general-purpose edge-feature guidance module to enhance vision transformers for plant disease identification," *Expert Systems with Applications*, vol. 237, p. 121638, 2024.

[20] Q. Hou, Z. Wang, F. Tan, Y. Zhao, H. Zheng, and W. Zhang, "RISTDnet: Robust infrared small target detection network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.

[21] Y. Dai, Y. Wu, F. Zhou, and K. Barnard, "Attentional local contrast networks for infrared small target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9813–9824, 2021.

[22] H. Duan, X. Xu, Y. Deng, and Z. Zeng, "Unmanned aerial vehicle recognition of maritime small-target based on biological eagle-eye vision adaptation mechanism," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 3368–3382, 2021.

[23] W. Liu, G. Ren, R. Yu, S. Guo, J. Zhu, and L. Zhang, "Image-adaptive YOLO for object detection in adverse weather conditions," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36,

no. 2, pp. 1792–1800, 2022.

[24] Q. Qin, K. Chang, M. Huang, and G. Li, "DENet: Detection-driven enhancement network for object detection under adverse weather conditions," in *Proceedings of the Asian Conference on Computer Vision*, pp. 2813–2829, 2022.

[25] Z. Cui, G. Qi, L. Gu, S. You, Z. Zhang, and T. Harada, "Multitask AET with orthogonal tangent regularity for dark object detection," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2553–2562, 2021.

[26] M. Li, H. Li, X. Ding, L. Wang, X. Wang, and F. Chen, "The detection of pine wilt disease: A literature review," *International Journal of Molecular Sciences*, vol. 23, no. 18, p. 10797, 2022.

[27] A. Coluccia, A. Fascista, and G. Ricci, "A KNN-based radar detector for coherent targets in non-Gaussian noise," *IEEE Signal Processing Letters*, vol. 28, pp. 778–782, 2021.

[28] X. Yin, Z. Yu, Z. Fei, W. Lv, and X. Gao, "PE-YOLO: Pyramid enhancement network for dark object detection," in *International Conference on Artificial Neural Networks*, pp. 163–174, 2023.

[29] Y. Deng, Z. Ren, Y. Kong, F. Bao, and Q. Dai, 'A hierarchical fused fuzzy deep neural network for data classification," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 4, pp. 1006–1012, 2016.

[30] Z. Zhou, Y. Hu, X. Yang, and J. Yang, "YOLO-based marine organism detection using two-terminal attention mechanism and difficult-sample resampling," *Applied Soft Computing*, vol. 153, p. 111291, 2024.

[31] C. Wang, J. Pei, X. Liu, Y. Huang, and J. Yang, "A deep deformable residual learning network for SAR image segmentation," in *2021 IEEE Radar Conference (RadarConf21)*, pp. 1–5, 2021.

[32] C. Ren, X. He, L. Qing, Y. Wu, and Y. Pu, "Remote sensing image recovery via enhanced residual learning and dual-luminance scheme," *Knowledge-Based Systems*, vol. 222, p. 107013, 2021.

[33] G. U. Nneji, J. Cai, D. Jianhua, I. A. Chikwendu, A. Oluwasanmi, E. C. James, and G. T. Mgbejime, "Enhancing low quality in radiograph datasets using wavelet transform convolutional neural network and generative adversarial network for COVID-19 identification," in *2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, pp. 146–151, 2021.

[34] H. Zhu, H. Wang, C.-T. Lam, L. Hu, B. K. Ng, and K. Fang, "Rapid APT detection in resource-constrained IoT devices using global vision federated learning (GV-FL)," in *International Conference on Neural Information Processing*, pp. 568–581, 2023.

[35] J. Chen, X. Yu, Q. Li, W. Wang, and B.-G. He, "LAG-YOLO: Efficient road damage detector via lightweight attention ghost module," *Journal of Intelligent Construction*, vol. 2, no. 1, 2024.

[36] K. Fang, T. Wang, X. Yuan, C. Miao, Y. Pan, and J. Li, "Detection of weak electromagnetic interference attacks based on fingerprint in IIoT systems," *Future Generation Computer Systems*, vol. 126, pp. 295–304, 2022.

[37] P. V. de Campos Souza, "Fuzzy neural networks and neuro-fuzzy networks: A review the main techniques and applications used in the literature," *Applied Soft Computing*, vol. 92, p. 106275, 2020.

[38] K. Zheng, Q. Zhang, Y. Hu, and B. Wu, "Design of fuzzy system-fuzzy neural network-backstepping control for complex robot system," *Information Sciences*, vol. 546, pp. 1230–1255, 2021.

[39] H. Fu, W. Liu, Y. Liu, Z. Cao, and H. Lu, "SIERRA: A robust bilateral feature upsampler for dense prediction," *Computer Vision and Image Understanding*, vol. 235, p. 103762, 2023.

[40] J. Gu, H. Cai, C. Dong, R. Zhang, Y. Zhang, W. Yang, and C. Yuan, "Super-resolution by predicting offsets: An ultra-efficient super-resolution network for rasterized images," in *European Conference on Computer Vision*, pp. 583–598, 2022.

[41] W. Liu, H. Lu, H. Fu, and Z. Cao, "Learning to Upsample by Learning to Sample," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6027–6037, 2023.

[42] S. K. Pandey, A. Shukla, S. Bhatia, T. R. Gadekallu, A. Kumar, A. Mashat, M. A. Shah, and R. M. Janghel, "Detection of arrhythmia heartbeats from ECG signal using wavelet transform-based CNN model," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, p. 80, 2023.

[43] K. Mao, R. Jin, L. Ying, X. Yao, G. Dai, K. Fang, "SC-YOLO: Provide Application-Level Recognition and Perception Capabilities for Smart City Industrial Cyber-Physical System," *IEEE Systems Journal*, 2023. DOI: 10.1109/JSYST.2023.3292113.

[44] H. Feng, Q. Li, W. Wang, A. K. Bashir, A. Kumar Singh, J. Xu, and K. Fang, 'Security of target recognition for UAV forestry remote sensing based on multi-source data fusion transformer framework," *Information Fusion*, vol. 112, pp. 102555, 2024.

[45] Z. Wei, Y. Wang, L. Sun, A. V. Vasilakos, and L. Wang, 'ClassLIE: Structure-and-Illumination-Adaptive Classification for Low-Light Image Enhancement," *IEEE Transactions on Artificial Intelligence*, 2024. DOI: 10.1109/TAI.2024.3405405.

[46] M. Xiao, L. Zhang, W. Shi, J. Liu, W. He, and Z. Jiang, "A visualization method based on the Grad-CAM for medical image segmentation model," in *2021 International Conference on Electronic Information Engineering and Computer Science (EIECS)*, pp. 242–247, 2021.

[47] J. Chu, J. Cai, L. Li, Y. Fan, and B. Su, "Bilinear feature fusion convolutional neural network for distributed tactile pressure recognition and understanding via visualization," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 6, pp. 6391–6400, 2021.