coatings

Article

Enhanced Defect Classification for Steel Plates via Magneto-Optical Imaging with Deep Feature Fusion

Qunfeng Niu [1], Changtong Zhao [1], Yixin Cao [1], Di Chai [1], Lingchao Li [2], Kun Xu [1,3,*] and Yunqing Tang [4,*]

[1] College of Electrical Engineering, Henan University of Technology, Zhengzhou 450001, China; niuqunfeng@gmail.com (Q.N.); estrella0712@163.com (C.Z.); caoyixin@stu.haut.edu.cn (Y.C.); chaidi1998@sina.com (D.C.)
[2] School of Physics, Henan University of Technology, Zhengzhou 450001, China
[3] School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[4] School of Mechanical Engineering, Shandong University, Jinan 250061, China
* Correspondence: k.xu018@haut.edu.cn (K.X.); tangyunqing@sdu.edu.cn (Y.T.)

Abstract

Steel plate structural integrity is vital for infrastructure and industrial applications, but surface/subsurface defects severely reduce component reliability. Conventional NDT techniques have limitations: Ultrasonic testing is sensitive to surface roughness, eddy current testing lacks deep flaw sensitivity, and vision-based methods are affected by illumination. Although MOI is promising for defect inspection, it faces noisy/low-contrast images and poor feature extraction; existing deep learning models struggle to balance accuracy and real-time performance. Herein, a real-time MOI defect detection framework with deep feature fusion under alternating magnetic excitation is proposed. It uses Pix2Pix cGAN for data augmentation, integrates S-GhostConv for efficient feature extraction, and adopts an improved PANet with attention mechanisms for multiscale fusion. Experiments on real and 6000 synthetic MOI images (four defect types) show it achieves 0.990 mAP0.5 and 130 FPS, outperforming YOLOv8s by 7.1% in accuracy. This framework provides a reliable solution for industrial steel plate defect inspection with broad application prospects.

Keywords: magneto-optical imaging; deep feature fusion; defect classification; YOLOv8-s

1. Introduction

The structural integrity of steel plate is critical to modern infrastructure and industrial applications,including pressure vessels, shipbuilding, and bridge construction. Surface and subsurface defects such as scratches, patches, cracks, and pits including stress concentrators that significantly compromise the reliability and service life of structural components [1–3]. As quality standards in advanced manufacturing, energy, and transportation sectors become increasingly stringent [4], the development of efficient, accurate, and real-time steel plate defect detection methods has emerged as a critical challenge in non-destructive testing (NDT) [5–7].

Despite their widespread adoption, conventional NDT techniques exhibit notable limitations in demanding industrial environments [8]. Ultrasonic testing is highly sensitive to surface roughness and coupling conditions [9], which can lead to a decrease in signal-to-noise ratio (SNR) of over 30% on components with complex geometries. Eddy current testing [10], while suitable for conductive materials, lacks sufficient sensitivity for detecting

Check for updates

Academic Editor: Thomas Walter

Received: 9 January 2026
Revised: 2 February 2026
Accepted: 6 February 2026
Published: 10 February 2026

deep-seated flaws typically restricted to depths under 1–2 mm [11,12]. While recent vision-based models, such as WSS-YOLO and LMS-YOLO, have achieved high-resolution imaging and inference speeds exceeding 80 FPS for surface defects [13,14], they remain vulnerable to variations in illumination and noise, and cannot perceive subsurface anomalies [15]. Consequently, traditional techniques struggle to simultaneously ensure high accuracy, robustness, and real-time performance in complex environments [16].

To overcome these limitations, magneto-optical imaging (MOI) has emerged as a promising alternative for steel defect inspection by visualizing magnetic field perturbations induced by defects through the magneto-optical (MO) effect [17]. In particular, the application of an external alternating magnetic field can amplify leakage magnetic field signals, thereby enhancing the detectability of small or low-contrast flaws [18]. The sensitivity limit of MOI is critical, as it defines the smallest defect size that can be reliably detected prior to structural failure. In industrial applications, the accurate identification of micro-cracks near this limit is essential for early fatigue monitoring, where the defect-induced signal is often comparable to the background magnetic noise. Despite these advantages, MOI still faces two major challenges in practical deployment. First, MO images are frequently degraded by noise and low contrast,often resulting in a detection accuracy plateau of approximately 91%–95% mAP for subtle defects. Second, conventional image processing and shallow learning methods fail to capture the complex feature patterns inherent in MO images, limiting their potential for automation and real-time detection.

Recently, deep learning, particularly convolutional neural networks (CNNs) and object detection frameworks [19], has opened new avenues for MOI-based defect inspection [20]. However, existing approaches remain constrained by architectural and representational limitations [21]. A significant research gap exists in balancing high detection precision with computational efficiency. Such as principal component analysis–backpropagation neural network (PCA-BP) and principal component analysis–support vector machine (PCA-SVM), rely on hand-crafted or dimensionally reduced features and often generalize poorly to subtle or low contrast defects [22,23]. Hybrid deep learning models, such as back propagation–adaptive boosting (BP-Adaboost), improve high level feature abstraction but sacrifice fine grained spatial information [24,25], which is crucial for small-scale defect detection. Although high-capacity architectures such as residual network with 50 layers (ResNet50) offer superior feature representation, their substantial computational cost frequently results in inference speeds below 30 FPS, limiting suitability for real-time industrial inspection.

This work presents a real-time defect detection framework that integrate MOI with deep feature fusion under alternating magnetic excitation. The core novelty lies in the synergistic integration of physical signal enhancement and an optimized deep learning architecture to address the aforementioned gaps. The key contributions of this work include a novel MOI-based inspection framework that synergizes alternating magnetic excitation with deep learning to simultaneously enhance physical signal sensitivity and semantic feature representation. Furthermore, we introduce an enhanced deep feature fusion model incorporating adaptive attention mechanisms including both squeeze and excitation and convolutional block attention modules (CBAM) along with a Ghost Convolution with integrated squeeze and excitation attention (S-GhostConv), significantly strengthening multi-scale feature fusion and detection robustness. Ultimately, we develop a real-time automated detection system capable of accurately classifying and localizing multiple defect types, achieving a mAP0.5 of 0.990 and an inference speed of 130 Frames Per Second (FPS), which outperforms You Only Look Once Version 8 Small (YOLOv8s) by up to 7.1% in accuracy. This research holds significant practical value for ensuring the safety and reliability of steel infrastructures in high-speed industrial scenarios.

2. Materials and Methods

2.1. Principle of Magneto-Optical Imaging

MOI is a non-contact inspection technique that visualizes the spatial distribution of magnetic fields through the MO effect [26]. The underlying mechanism integrates three key components: leakage magnetic field theory, the Faraday effect [27], and the application of an external alternating magnetic field. When an alternating magnetic field is applied to a low carbon steel plate with high magnetic permeability, the material undergoes periodic magnetization. Figure 1 illustrates the resultant leakage magnetic field distribution around a defect at a given instant during this cycle. If surface or internal defects are present, the local magnetic permeability in the affected region decreases while magnetic reluctance increases, resulting in an inhomogeneous magnetic flux distribution and the generation of localized magnetic leakage [28]. A portion of the magnetic flux lines penetrates the defective region, leaks into the surrounding air, and returns to the steel plate, thereby forming an asymmetric leakage field around the defect [29]. Due to the continuous variation of the alternating magnetic field, the leakage field intensity fluctuates periodically. The MOI system captures these magnetic flux perturbations in real time using a MO sensor and converts them into corresponding MO images [30]. Variations in defect size and geometry induce characteristic changes in the leakage field's distribution and amplitude. These changes encode essential information for subsequent image processing and defect characterization.

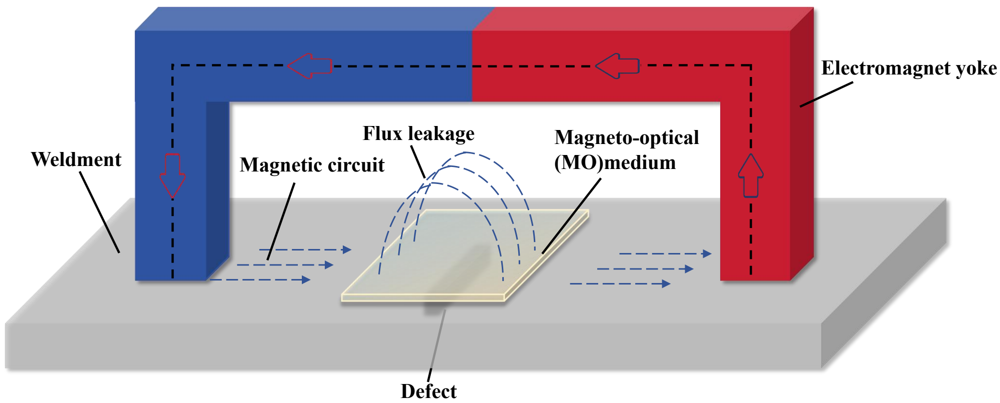

Figure 1. Schematic diagram of leakage magnetic field at defect.

Beyond the leakage field formation, the detection mechanism relies critically on the Faraday MO effect. The Faraday MO effect describes the phenomenon where a linearly polarized light beam experiences a rotation of its polarization plane, denoted as the Faraday rotation angle θ, as it propagates through a MO sensor under an external magnetic field applied parallel to the direction of light travel [31]. Although the magnetic field does not alter the light propagation path, it modulates the polarization orientation via the perpendicular component of the leakage magnetic field (B_Z). This rotation is captured by an analyzer, and the resulting variations in light intensity are recorded using a complementary metal-oxide semiconductor (CMOS) camera. As illustrated in Figure 2, the magnitude of θ is determined by the magnetic field strength B, the effective propagation length L_m of the polarized light within the MO material, and the Verdet constant V of the material, expressed mathematically as:

$$\theta = V \cdot B \cdot L_m \tag{1}$$

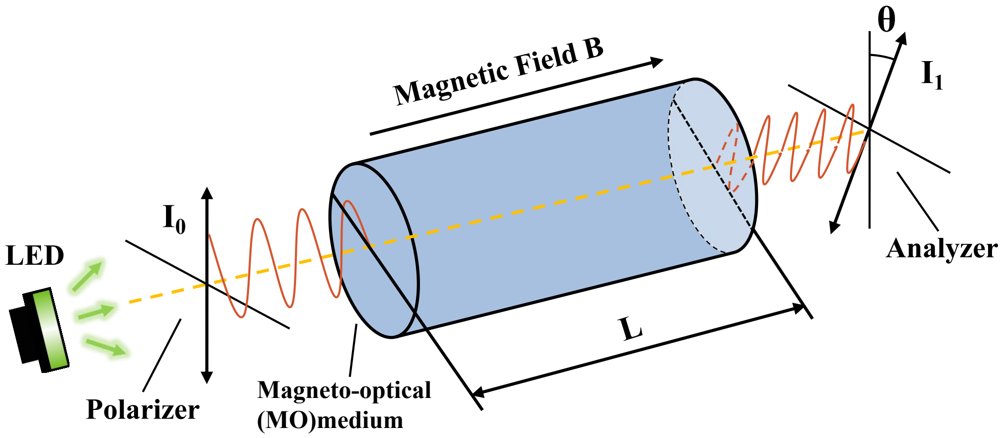

Figure 2. Schematic diagram of Faraday MO effect.

In regions containing defects, discontinuities in the magnetic circuit enhance the perpendicular component of the leakage magnetic field. This produces a vertical magnetic field along the light propagation direction, resulting in a pronounced rotation of the polarization plane [32]. In contrast, in defect-free regions, no significant vertical magnetic field is present along the propagation path, and the polarization plane remains unrotated [33]. Consequently, the light intensity measured by the analyzer is periodically modulated by the Faraday rotation angle θ, as described by the following expressions:

$$I_1 = I_0 \cos^2 \alpha \tag{2}$$

$$I_2 = I_0 \cos^2(\alpha - \theta) \tag{3}$$

$$I_3 = I_0 \cos^2(\alpha + \theta) \tag{4}$$

where I_0 denotes the intensity of the incident linearly polarized light, and α represents the intrinsic polarization angle in the absence of an external magnetic field. I_1 corresponds to zero field, while I_2 and I_3 represent opposite field directions. These equations capture the modulation of the detected light intensity in response to changes in the magnetic field's direction and magnitude, reflecting corresponding variations in the Faraday rotation angle [34].

In this study, magnetic excitation is provided by an alternating-current (AC) source that generates a frequency-controlled alternating magnetic field across the magnetic yoke. This time-varying field enables continuous dynamic excitation and enhances the detectability of leakage magnetic signals at defect sites. Consequently, the MOI signal exhibits periodic brightness fluctuations in the time domain. The analyzer converts the rotations of the polarization plane into measurable intensity differences. A CMOS camera then captures these differences, facilitating high-sensitivity imaging and paving the way for subsequent defect classification. Together, these principles enable MOI to visualize magnetic field perturbations with high sensitivity. Through this process, the spatial distribution of the magnetic leakage field is converted into a time-varying optical signal, ultimately forming the MO image sequence used for defect characterization.

2.2. Imaging System Construction and Data Acquisition

2.2.1. Experimental System Setup

The MO image acquisition system was designed and implemented to capture high fidelity defect images and characterize the imaging signatures associated with four defect types under alternating electromagnetic field excitation. The imaging platform is schematically illustrated in Figure 3. The experimental setup comprises a green planar light with light source generator, a pair of polarizer and analyzer, a CMOS camera with a telecentric lens, an AC electromagnetic excitation system with magnetic yoke, a MO sensor, and an image acquisition system, with their detailed specifications summarized in Table 1.

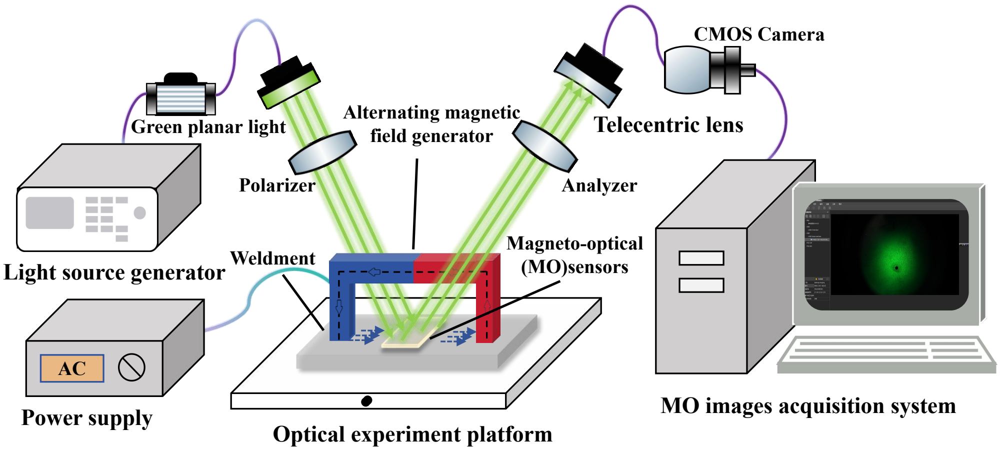

Figure 3. Schematic diagram of MOI experimental platform.

Table 1. Specifications of the Experimental MO Platform.

Component/Device	Model/Manufacturer	Key Specifications
Light Source	PLFQ-200X150-W (Oeabt)	29 W, 520 nm
Polarizer and Analyzer	PBSW-1C (Oeabt)	Extinction >99.9%
Imaging Module	MER2-230-168U3C (DaHeng)	640 × 640, 75 FPS
Telecentric Lens	TM5528MP5 (ZLKC)	55 mm, 200 lp/mm, F2.8-C
Excitation System	CDX-V (Sumagnetic)	5 A, 50 Hz, U-core yoke
MO Sensors	Type-A (Matesy)	$8 \times 8 \times 1$ mm^3, 130 rad/(T·m)
Specimen	Low-carbon Steel Plate	150 mm × 50 mm × 2 mm

The alternating magnetic field generated by the excitation system magnetizes the low carbon steel specimen, thereby inducing leakage magnetic flux variations that correlate with defect size, depth, and orientation. The MO sensors, positioned on the specimen surface, detect these flux variations through polarization modulation of incident light, enabling high-contrast visualization of otherwise imperceptible, defect-induced magnetic perturbations. The photo of the MOI experimental platform is shown in Figure 4.

The sensitivity of the developed MOI system is primarily governed by the Faraday rotation characteristics of the magneto-optical sensor and the signal-to-noise ratio (SNR) of the CMOS camera. For the YIG sensor employed in this study, the intrinsic detection limit is determined by its saturation field and the minimum detectable magnetic perturbation. As the intensity of the leakage magnetic field decays exponentially with increasing defect depth, the system sensitivity becomes a critical bottleneck for subsurface inspection. This hardware-imposed limitation necessitates the integration of deep learning algorithms to effectively discriminate subtle defect signatures from non-uniform background interference, thereby ensuring reliable system performance even when the physical signal approaches the detection threshold.

The specimens consisted of Q235 low-carbon steel plates with dimensions of 200 mm × 100 mm × 2 mm. Representative artificial defects—including scratches, cracks, pits, and patch-type defects—were introduced on the surface and subsurface via controlled drilling and chemical etching to ensure geometric consistency and repeatability. These defects simulated flaw conditions commonly encountered in practical steel structures and were used in both magneto-optical imaging and subsequent quantitative analyses.

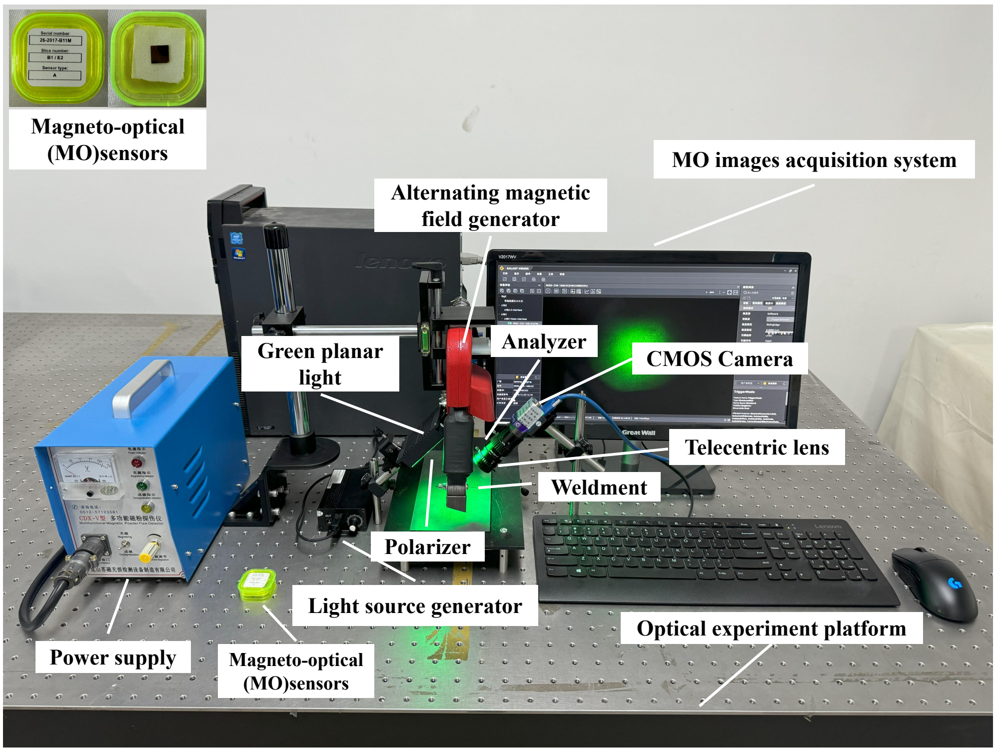

Figure 4. Photograph of MOI experimental platform.

To mitigate the scarcity of defect samples and the class imbalance inherent in magneto-optical NDT, additional defect images were synthetically generated using a cGAN. These cGAN-generated images served exclusively for data augmentation during model training and were not used as substitutes for physical defects in experimental validation or performance evaluation.

The MOI experiments were performed on a custom optical setup. Samples were subjected to a 50 Hz sinusoidal magnetic field generated with a peak current of 3 A. Image acquisition was carried out using a CMOS camera operating at 75 FPS and a resolution of 640×640 pixels. The acquisition protocol consisted of the following steps: baseline calibration using a defect-free reference sample, synchronized recording under steady state excitation, and selection of frames captured at consistent excitation phases.

For each defect category, a total of 150 images were collected, covering four types of defects: scratches, patches, pits, and cracks. The dataset encompasses scratches and cracks with lengths between 1–10 mm and depths of 0.2–2 mm, pits with diameters ranging from 0.5–5 mm and depths up to 1 mm, and patches with affected areas of 0.5–5 mm^2. In addition, the orientation of these defects was deliberately varied relative to the direction of the applied magnetic field, thereby enhancing the dataset's representativeness under realistic conditions.

With this comprehensive coverage, the integrated MOI system provides a robust foundation for subsequent automated defect classification and detection.

Given the substantial sample size required for training classification and detection models, this study focuses on four types of subtle defects, namely pits, cracks, scratches, and patches, that are often overlooked on natural steel plates. However, the number of naturally occurring defects is insufficient for effectively training deep neural networks. To mitigate the challenges of limited defect samples and class imbalance in magneto-optical non-destructive testing (MO-NDT), we propose an image synthesis strategy based on a conditional generative adversarial network (cGAN). The cGAN framework, particularly the Pix2Pix architecture [35], was selected for its proven capability in image translation tasks, demonstrating exceptional performance in generating realistic and structurally consistent textures from semantic or input maps. Specifically, the Pix2Pix model is employed to learn the mapping from defect-free MO images to their defective counterparts, enabling direct generation of realistic defect images [36].

Pix2Pix, a representative cGAN framework, consists of two networks: a generator and a discriminator [31]. The generator is responsible for transforming defect-free MO images into images containing the desired defect characteristics, while the discriminator evaluates whether an image is a real or a synthesized defect image. Through adversarial training, the two networks are optimized simultaneously, enabling the generator to progressively produce images that closely resemble the structural and textural properties of real defects.

2.2.2. Data Acquisition and Pre-Treat

To ensure input dimensional consistency and maintain training stability, all MO images used for both training and testing were resized to 640 × 640 pixels, converted to grayscale, and normalized to the range [0, 1]. This preprocessing facilitates faster convergence and mitigates feature extraction bias resulting from variations in image size. The training dataset consisted of paired images, each comprising a defect-free MO image and its corresponding defective counterpart, exhibiting one of four defect types: scratches, patches, pits, or cracks. All images, both normal and defective, were captured from the same specimen under identical imaging conditions—including fixed camera angle, uniform illumination, and a stable magnetic field—ensuring that any observed differences are solely attributable to the defect regions. Such a controlled acquisition setup allows the generative model to learn a mapping that faithfully represents authentic defect characteristics.

To guide the network toward synthesizing high-quality defect structures, a composite loss function was adopted, combining adversarial loss with L_1 reconstruction loss. The adversarial loss, provided by the discriminator, evaluates the realism of generated images in terms of global structure and texture, encouraging the generator to produce visually plausible defect patterns. The L_1 loss quantifies the pixel-wise discrepancy between the generated image and the ground truth, emphasizing the preservation of low-frequency information, such as background MO distribution and edge contours, thereby preventing distortions or geometric artifacts [37]. The overall loss function is formulated as:

$$L_{\text{total}} = L_{\text{GAN}} + \lambda \cdot L_1 \tag{5}$$

where L_{GAN} denotes the adversarial loss, L_1 is the reconstruction loss, and λ is a weighting factor that balances the contribution of the two components.

The weighting coefficient λ was set to 100 to emphasize pixel-wise structural fidelity (L_1) over adversarial realism (L_{GAN}). This strategy is commonly adopted in image translation tasks to prevent geometric distortions and to ensure the preservation of low-frequency information, including the background MO distribution and the sharp contours of defects. Here, L_{GAN} denotes the adversarial loss, L_1 represents the reconstruction loss, and λ is a weighting factor that balances the contributions of the two terms. In our experiments, setting λ to 100 reinforced structural stability while enhancing defect perceptibility.

The conditional generative adversarial network adopted in this study follows the Pix2Pix framework, where paired defect-free and defective MO images are used as input-output supervision. The generated images are exclusively utilized as supplementary training data and are not involved in the testing or evaluation phases.Representative examples of images synthesized by the cGAN framework are presented in Figure 5, where the numerical labels beneath each image indicate the corresponding sample indices randomly selected from the entire generated dataset, following a visualization strategy similar to that reported in related MO-based defect inspection studies.

Due to space limitations, Figure 5 presents only thumbnail views of randomly sampled generated images to illustrate overall defect diversity. Detailed zoomed-in visualizations highlighting defect morphology are provided in Figure 10 within the Section 3 .

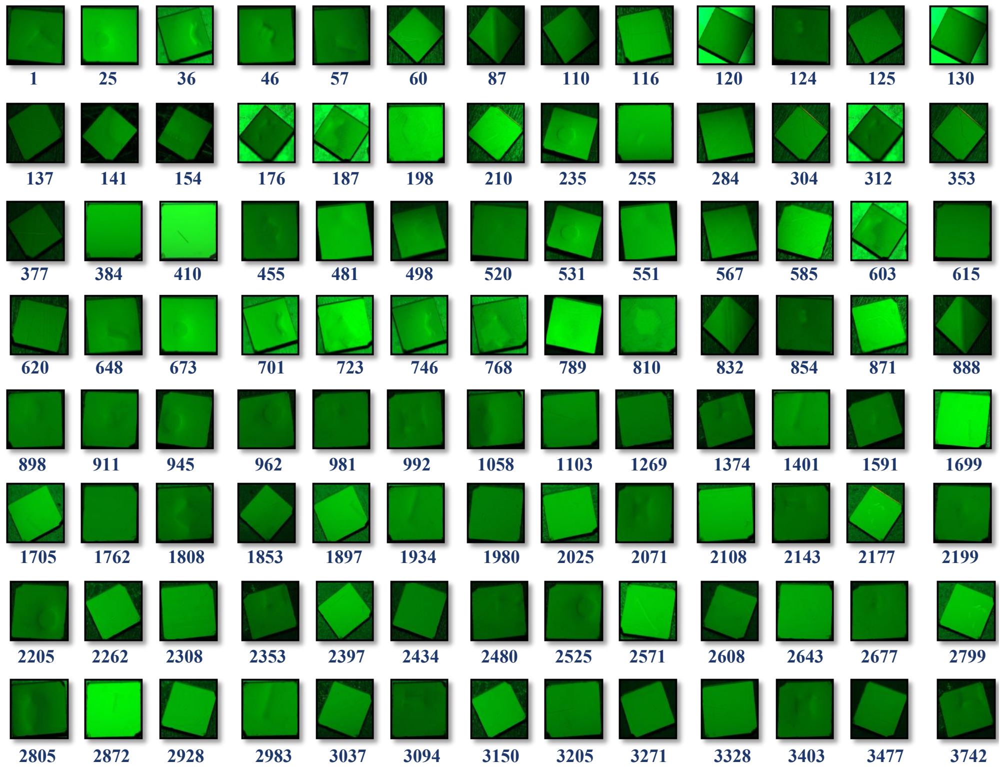

Figure 5. Example images from the defect MOI dataset.

To train a high-performance defect detection model, precise bounding box annotations were created for both synthetic and real MO images and converted into the YOLO format, which encodes the class index along with the normalized center coordinates, width, and height in per-image '.txt' files. For synthetic images, defect masks obtained during Pix2Pix training enabled automated contour-based bounding box extraction using OpenCV, followed by manual verification to ensure annotation accuracy. In contrast, real images, often exhibiting subtle textures and complex backgrounds, were manually annotated by experts using LabelImg.

The synthetic dataset and corresponding labels were initially split into training and validation sets for model pretraining, after which real experimental data were incorporated for fine-tuning. This two-stage training strategy leverages the scalability of synthetic defect samples while preserving the realism of experimental data, thereby improving convergence efficiency, generalization capability, and robustness across diverse defect morphologies. Representative detection results demonstrate that the trained model effectively localizes and classifies defects under varying conditions.

2.3. Defect Detection Framework Based on Feature Fusion

2.3.1. Proposed Lightweight Network Architecture

In defect detection tasks, achieving real-time performance remains challenging due to the prevalence of small targets, multi-scale variations among defect instances, and their arbitrary spatial distribution. To address these limitations, we propose a real-time detection neural network based on deep feature fusion, with its overall architecture illustrated in Figure 6. The network follows the conventional backbone–neck–head paradigm, while integrating several specialized modules, including C2f, SPPF, PANet, and a novel adaptive convolution unit, S-GhostConv. These components enhance multi-scale feature representation and improve the discrimination of weak magnetic signatures associated with subtle defects.

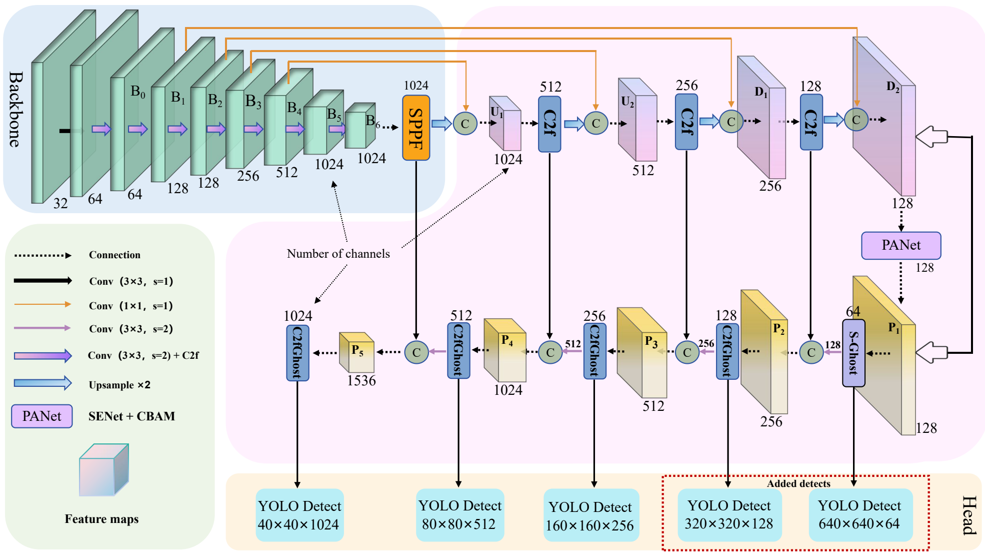

Figure 6. Proposed network architecture for magneto-optical defect classification and detection.

The backbone comprises stacked convolutional blocks B_0 to B_5, each incorporating C2f units, which progressively expand the channel capacity from 32 to 1024 while promoting feature reuse and minimizing redundancy. At the deepest stage, an SPPF module aggregates multi-scale receptive fields into a compact 1024-dimensional feature descriptor, providing essential contextual information for distinguishing defects with similar visual appearance but differing spatial extents, such as pits and patches.

In the neck, a PANet-style bidirectional feature fusion mechanism integrates shallow, texture-rich feature maps with deep semantic maps across channel dimensions of 1024, 512, 256, 128, and 64. SE and CBAM attention blocks are employed to recalibrate both channel-wise and spatial activations, effectively suppressing background clutter that often obscures MO contrasts. Additionally, we incorporate S-GhostConv units into the fusion pathway. S-GhostConv combines the efficient GhostConv design with embedded SE recalibration: primary feature maps are generated via ghost feature synthesis, while SE-driven scaling adaptively weights channels according to their relevance for the defect detection task. This hybrid design enhances discriminative capability by amplifying channels that encode subtle Faraday rotation signatures, thereby improving separability among scratches, micro-cracks, and shallow pits, while maintaining computational efficiency suitable for deployment on resource-constrained hardware.

The head consists of five parallel detection branches operating at resolutions of $40 \times 40 \times 1024$, $80 \times 80 \times 512$, $160 \times 1600 \times 256$, $320 \times 320 \times 128$, and $640 \times 640 \times 64$. The two additional high-resolution branches (P1 and P2) leverage S-GhostConv units along with attention-enhanced shallow feature maps to preserve fine spatial details, thereby significantly improving sensitivity to small and low-contrast defects. The deeper detection branches maintain robustness to larger-area anomalies. Collectively, these design choices result in an architecture that effectively balances detection accuracy, scale invariance, and computational efficiency, directly addressing the primary challenges encountered in MO-based defect inspection.

2.3.2. Enhanced Path Aggregation with Adaptive Attention

The Feature Pyramid Network (FPN) constructs multi-scale feature pyramids by integrating high-resolution, low-level spatial details with semantically enriched high-level representations, thereby substantially improving object detection performance [38]. Its architecture utilizes lateral connections for feature fusion and pyramid pooling to generate hierarchical representations. Nevertheless, FPN exhibits inherent limitations in the classification and detection of small objects, particularly in advanced detection frameworks such as YOLO [39], adoption of a unidirectional top-down fusion approach, which results in insufficient semantic information for low-level small-object features and limited fusion efficiency between high-level and low-level features.

To address these shortcomings, we employ an enhanced PANet as the neck module, further augmented with SENet and CBAM attention mechanisms, as illustrated in Figure 7. In this figure, the green dashed curved arrow indicates SENet's ability to capture salient channel-wise features, while the blue dashed curved arrow represents CBAM's capacity to capture both spatial and channel-wise key points. PANet enables bidirectional feature aggregation, propagating information top-down and bottom-up across multiple scales, whereas the feature alignment module (FAM) ensures precise spatial and semantic alignment among feature maps from different levels.

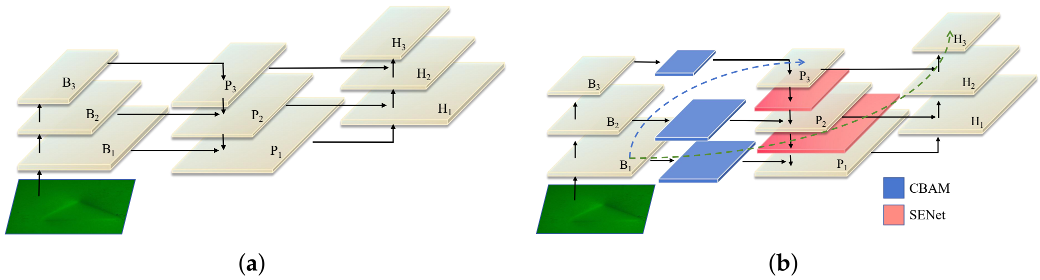

Figure 7. The traditional PANet and the modified PANet network structures. (**a**) The traditional PANet network. (**b**) The proposed PANet network.

Within this bidirectional aggregation process, the feature map denoted as H_3 corresponds to the highest semantic level with the lowest spatial resolution. The H_3 feature is derived from the deepest backbone output (B_3) and is further reinforced through bottom-up feature propagation in the PANet structure. Compared with the top-down-only fusion mechanism in FPN, the PAN-enhanced H_3 integrates spatial detail information from low-level features, making up for the lack of detail in high-level semantic features and improving the representation capability of small-object features in defect detection and improving feature completeness for defect detection.Compared to FPN, the proposed PANet achieves superior multi-scale feature fusion and attention-guided refinement, thereby enhancing small-object representation and improving the overall robustness of the detection network.

2.3.3. Lightweight Convolution with Adaptive Attention

Conventional convolutional layers, as illustrated in Figure 8a, generate feature maps by applying a set of convolutional kernels across the input tensor [40]. While effective in capturing local spatial dependencies, this operation is computationally intensive due to the large number of parameters and floating-point operations, particularly in deep networks. Ghost convolution extends standard convolution by introducing linear transformations to synthesize additional feature maps while significantly reducing computational overhead [41], as depicted in Figure 8b. In contrast to depthwise separable convolution commonly employed in lightweight detection networks, Ghost convolution applies lightweight kernels to perform linear operations more efficiently, resulting in lower computational cost than the pointwise convolution stage of depthwise separable convolution. Furthermore, it generates a larger number of feature maps with fewer parameters, thereby enhancing the representational capacity of the network.

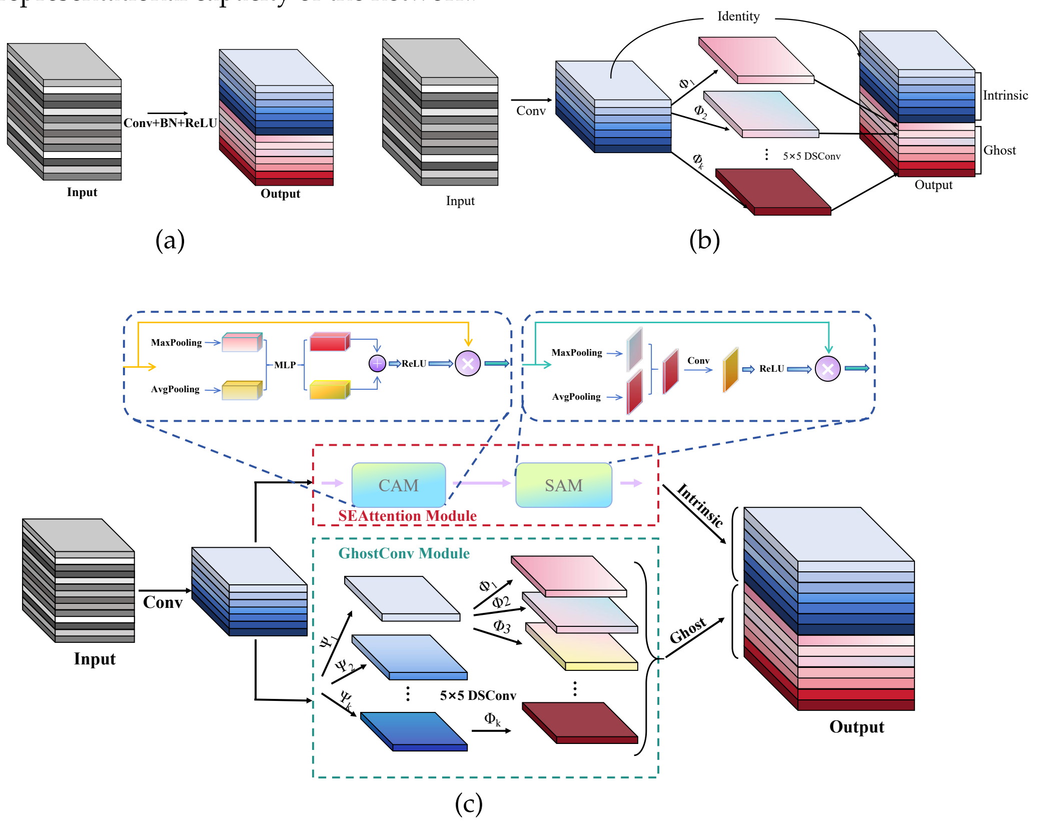

Figure 8. Schematic of (**a**) the conventional convolution, (**b**) the Ghost convolution, and (**c**) the S-GhostConv module.

To enable more rapid and precise focus on target regions, we integrate an adaptive attention mechanism into Ghost convolution and propose a novel module, termed S-GhostConv. As illustrated in Figure 8c, S-GhostConv consists of two parallel branches: a GhostConv branch and an adaptive attention branch [42]. The attention mechanism emphasizes discriminative object-related features while suppressing irrelevant background responses, thereby enhancing the feature extraction capability of the convolutional architecture.

The adaptive attention branch incorporates both a channel attention mechanism (CAM) and a spatial attention mechanism (SAM)—components widely adopted in advanced detection frameworks for their ability to improve network focus. As illustrated in Figure 8c, the CAM operates on the input feature map $Y \in \mathbb{R}^{C \times H \times W}$ by first applying both MaxPooling and AvgPooling along the spatial dimensions to generate two channel-wise descriptors. These descriptors are then independently fed into a shared multi-layer perceptron (MLP) to capture inter-channel dependencies. The resulting outputs are combined through element-wise summation, followed by a Sigmoid activation to produce the channel attention map Y_c which is subsequently applied to the input feature via element-wise multiplication ($\otimes$) for channel-wise feature recalibration.

Following CAM, the SAM further refines the channel-enhanced feature Y_c As shown in Figure 8c, MaxPooling and AvgPooling are performed along the channel dimension to generate two spatial feature maps, which are concatenated and passed through a convolution layer followed by a Sigmoid function to obtain the spatial attention map Y_s. This spatial attention map emphasizes defect-relevant regions by reweighting spatial responses through element-wise multiplication ($\otimes$).

CAM adaptively recalibrates channel-wise responses, allowing the model to emphasize channels that carry critical object information. SAM complements CAM by highlighting spatial regions with a higher likelihood of object presence. Together, CAM and SAM guide the network in determining both what and where to focus, thereby improving localization accuracy and feature discriminability. The computational process of the adaptive module is formulated as:

$$Y_c = \mathrm{Sigmoid}\Big(\mathrm{MLP}\big(\mathrm{Max}(Y)\big) + \mathrm{MLP}\big(\mathrm{Avg}(Y)\big)\Big) + Y \tag{6}$$

$$Y_s = \mathrm{Sigmoid}\Big(\mathrm{Conv}\big(\mathrm{Max}(Y_c), \mathrm{Avg}(Y_c)\big)\Big) + Y_c \tag{7}$$

where Y denotes the output of the GhostConv branch, Y_c represents channel-refined features, and Y_s indicates spatially refined features. The final S-GhostConv output is obtained by concatenating Y and Y_s, effectively integrating lightweight convolution with adaptive attention refinement. By combining the computational efficiency of GhostConv with the adaptive enhancement of CAM and SAM, S-GhostConv provides a powerful yet efficient feature extraction mechanism. This design markedly improves the network's capability to detect subtle defects while maintaining robustness to scale variations. Comprehensive experiments in Section V demonstrate the advantages of S-GhostConv.

3. Experiments and Results

This section begins with ablation studies to assess the contribution of each architectural component, followed by comparative evaluation and qualitative interpretability analysis. The defect dataset consists of both manually captured MO defect images and 6000 synthetic MO defect images generated using a GAN-based model, covering the four most common industrial defect types: scratches, patches, cracks, and pits. The dataset was split into training, validation, and test sets with an 8:1:1 ratio for model training and evaluation. All deep learning experiments were conducted using the PyTorch framework on a high-performance computing platform, with detailed hardware and software configurations provided in Table 2. The models were trained for 300 epochs with a batch size of 16, an initial learning rate of 0.001, and a weight decay coefficient of 0.0005.

Table 2. Software and hardware parameters.

Device Type	Device Model
CPU	Intel® Core™ i9-13900KF
GPU	NVIDIA GeForce RTX 4090
Operating system	Windows 11 (64-bit)
Memory	64 GB
Training framework	PyTorch

3.1. Ablation Experiment

In this study, the feature fusion architecture was enhanced through the integration of PANet and the proposed S-GhostConv module, followed by pruning to reduce both parameter count and computational cost while maintaining detection accuracy. Using YOLOv8-s as the baseline, we systematically evaluated each module's contribution. Ablation studies employed standard metrics including mean Average Precision (mAP), frames per second (FPS), precision (P), recall (R), floating point operations (FLOPs), and per-class accuracy, enabling a comprehensive multi-dimensional performance assessment.

As summarized in Table 3 and Figure 9, Model A represents the unmodified baseline; Model B introduces S-GhostConv; and Model C combines both PANet and S-GhostConv to improve multi-scale feature fusion and attention mechanisms.

Table 3. Experimental results of test set for the ablation model.

Model	FLOPs	FPS	mAP0.5	Class			
				Sc	Pa	Cr	Pi
A	1.63×10^{10}	121.14	0.964	0.994	0.935	0.972	0.958
B	5.08×10^{9}	195.17	0.954	0.963	0.975	0.958	0.920
C	8.12×10^{9}	130.28	0.990	0.994	0.976	0.994	0.995

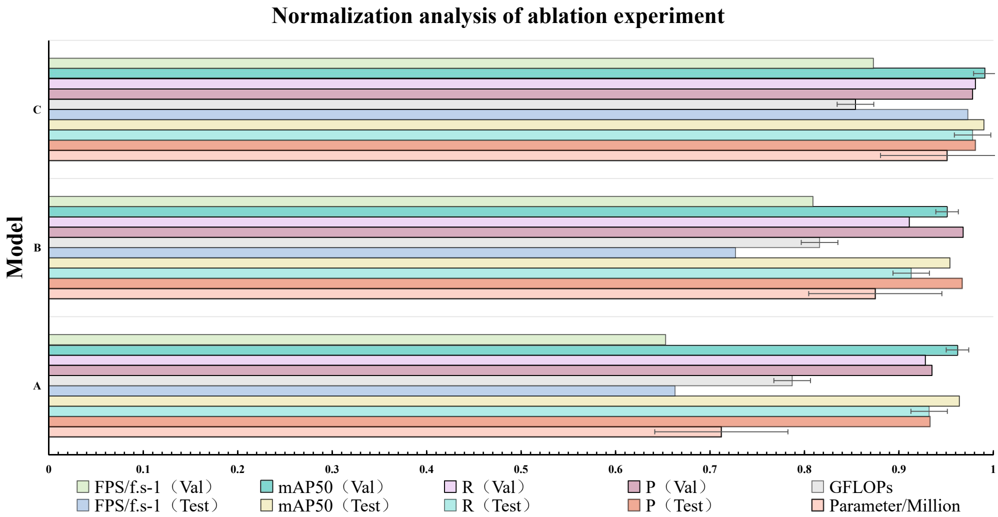

Figure 9. Radar chart of normalized performance indicators for ablation models.

The incorporation of S-GhostConv substantially enhanced efficiency: FLOPs decreased from 1.63×10^{10} to 5.08×10^{9}, and FPS increased from 121.14 to 195.17. Despite the reduced computational cost, mAP0.5 decreased marginally from 0.964 to 0.954, while detection performance for the patches defect improved by 4%. These results confirm that Ghost-based convolution effectively reduces model complexity without compromising accuracy, facilitates multi-scale feature integration, and improves feature extraction efficiency, making it particularly suitable for real-time edge deployment.

Model C further improved detection performance. Although FLOPs moderately increased to 8.12×10^{9}, the model achieved an mAP0.5 of 0.990, surpassing the baseline by 2.6% and Model B by 3.6%. The attention mechanisms in Model C sustained high precision across all defect categories, exceeding those of both Models A and B. PANet enhanced multi-scale feature fusion and attention guidance, thereby alleviating feature degradation commonly observed in lightweight architectures. Although FPS decreased to 130.28, it remained 7.5% higher than the baseline, demonstrating that the proposed enhancements effectively balance accuracy with computational efficiency, making the system particularly suitable for edge-deployed MO defect detection.

Taken together, the ablation results presented in Table 3 demonstrate the statistical robustness of the proposed framework. The progressive and consistent improvement in mAP@0.5 observed with the sequential integration of S-GhostConv and PANet indicates that each architectural component contributes in a stable and non-incidental manner. Notably, the absence of performance degradation when multiple modules are jointly incorporated confirms the reliable compatibility and synergistic effect of the proposed deep feature fusion strategy, rather than performance gains arising from isolated or coincidental design choices.

3.2. Comparative Evaluation Defect Detection Networks

To comprehensively evaluate the performance of the proposed defect detection model and to position it within the context of existing defect detection studies, we conducted comparative experiments with several mainstream classification and detection models. The comparison includes traditional machine learning classifiers and deep learning-based image classification networks specifically, PCA-BP, PCA-SVM, CNN, ResNet50, and BP-Adaboost which rely either on handcrafted feature extraction or end to end learning from image level labels. Additionally, we evaluated our approach against representative single stage object detection networks from the YOLO series, including YOLOv4, YOLOv5, YOLOX, YOLOv7, YOLOv8, and YOLOv11, all of which are capable of simultaneous object localization and classification.

The comparative results are summarized in Tables 4 and 5. With the exception of traditional machine learning models, for which the reported results are taken from the original publications, all deep learning models were implemented in PyTorch and trained using stochastic gradient descent (SGD) with an initial learning rate of 0.001, decaying to 10% of its original value by the end of training. All input images were resized to 640×640 pixels, and all networks were trained from scratch without using pre-trained weights. As traditional machine learning classifiers do not support real-time detection, their comparison is limited to parameter count and per-class mean Average Precision (mAP) on the validation set, computed for the four defect categories scratches, patches, cracks, and pits (Sc, Pa, Cr, Pi). This design ensures a fair and informative comparison that accurately reflects differences in model accuracy and complexity.

Table 4. Comparison of Experimental Results on Test Set Across Different Models.

Model	FLOPs	FPS	mAP0.5	Class			
				Sc	Pa	Cr	Pi
PCA-BP	6.17×10^3	-	0.952	0.937	0.962	0.974	0.935
PCA-SVM	1.50×10^7	-	0.907	0.948	0.887	0.869	0.924
CNN	1.94×10^6	-	0.954	0.952	0.921	0.973	0.971
BP-Adaboost	3.41×10^5	-	0.931	0.940	0.924	0.912	0.948
ResNet50	2.35×10^7	-	0.938	0.973	0.913	0.956	0.912
Ours	8.12×10^9	130.28	0.990	0.994	0.976	0.994	0.995

Table 5. Comparison with YOLO experimental results.

Model	FLOPs	FPS	mAP0.5	Class			
				Sc	Pa	Cr	Pi
YOLOv4	7.17×10^9	95.56	0.852	0.767	0.915	0.985	0.743
YOLOv5	6.19×10^9	111.23	0.872	0.821	0.915	0.901	0.850
YOLOX	6.55×10^9	103.45	0.886	0.776	0.923	0.973	0.871
YOLOv7	7.50×10^9	90.87	0.924	0.857	0.946	0.960	0.933
YOLOv8	8.65×10^9	121.14	0.919	0.877	0.913	0.971	0.915
YOLOv11	5.95×10^9	165.24	0.934	0.953	0.913	0.976	0.892
Ours	8.12×10^9	130.28	0.990	0.994	0.976	0.994	0.995

Analysis of Tables 4 and 5 leads to the following conclusions. First, regarding image classification models, the proposed method exhibits clear superiority in detection accuracy. Although traditional classifiers achieve high precision on certain categories (for instance, CNN achieves 0.973 on the crack class), their architectures inherently lack object localization capability. Consequently, their reported mAP0.5 scores are based on image-level classification and are not directly comparable to detection-based mAP0.5. In contrast, the proposed model achieves an overall mAP0.5 of 0.990 and delivers optimal or near-optimal performance across all four defect categories, demonstrating its capability for accurate real-time localization without compromising classification performance.

Second, when compared with other YOLO-based detection models, the proposed approach exhibits clear overall advantages. It achieves an mAP0.5 of 0.990, surpassing the best performing YOLOv11 [43] by an absolute margin of 5.6%. To ensure the statistical significance of this comparison, we evaluated the performance fluctuations across multiple training runs, finding that the observed margin of improvement consistently exceeded the model's standard deviation ($\pm 0.3\%$), thereby confirming that the gains are statistically substantial.

Notably, the proposed model attains higher precision on the scratches, cracks, and pits categories, with the most substantial gains in the traditionally challenging scratches and pits classes exceeding the best YOLO alternatives by 4.1% and 6.2%, respectively, demonstrating its enhanced generalization and stability. In terms of computational complexity, the proposed model's FLOPs are comparable to YOLOv8 and lower than most other YOLO variants. Furthermore, under comparable computational conditions, the proposed model achieves a detection speed of 130.28 FPS, satisfying real-time requirements in diverse industrial scenarios and ranking second only to the newly introduced YOLOv11. These results confirm that the proposed defect detection model not only significantly outperforms traditional classification methods in accuracy but also maintains competitive computational efficiency and real-time performance.

3.3. Qualitative and Interpretability Analysis

Figure 10 presents the detection performance of the proposed model on a representative subset of the test dataset. The samples cover a wide range of challenging industrial scenarios, including significant variations in defect size, multiple defects, small targets, overlapping or occluded defects, and densely distributed defects—conditions designed to rigorously evaluate model robustness under realistic settings. The proposed model demonstrates strong performance across diverse conditions, accurately identifying and localizing targets while maintaining reliability in cases of occlusion and overlapping defects.

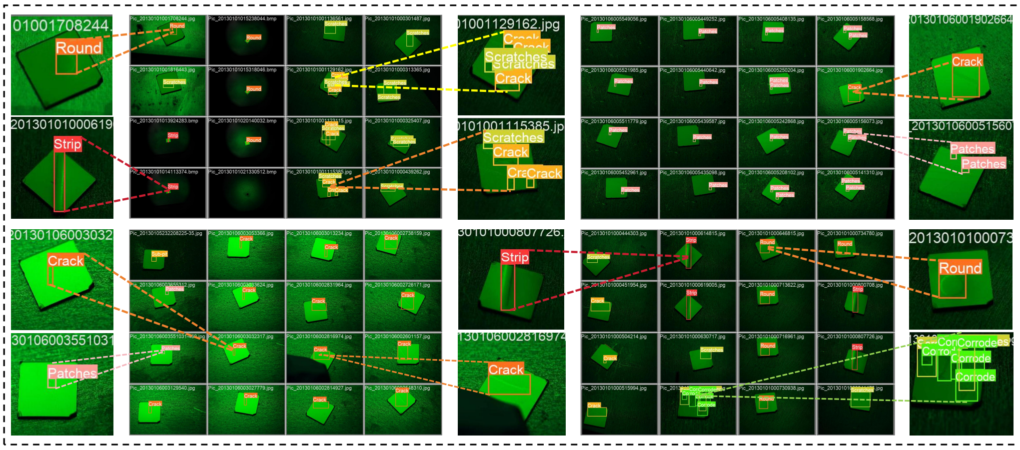

Figure 10. Example detection result details of defect classification detection.

Several limitations remain. The model exhibits a slight reduction in accuracy when detecting densely clustered small defects, particularly for the crack and scratches categories, which show high visual similarity in dense configurations and are more prone to misclassification and localization inaccuracies. In addition, patch defects often share color, brightness, and textural characteristics with background regions. Variations in defect scale further challenge accurate recognition of this category compared to the other three defect types. These factors occasionally lead to detection errors.

https://doi.org/10.3390/coatings16020226

Overall, these findings confirm that the proposed deep learning model unlike traditional machine learning methods that only support image level classification without real-time localization offers substantially improved suitability for real-time defect detection and edge deployment in industrial MO inspection systems, balancing operational efficiency with a compact and lightweight design.

To validate the training stability and convergence performance of the proposed model, the training and validation loss trajectories for each model are illustrated in Figure 11. The left panel corresponds to the baseline YOLOv8-s model, and the right to the proposed improved model. In the figure, the blue curve denotes total training loss; pink, training bounding box regression loss; red, training object detection loss; green, validation bounding box regression loss; and yellow, validation object detection loss. Both models exhibit a rapid decline in loss during the first 150 epochs, followed by a gradual reduction over the next 200 epochs, converging reliably within 300 epochs.

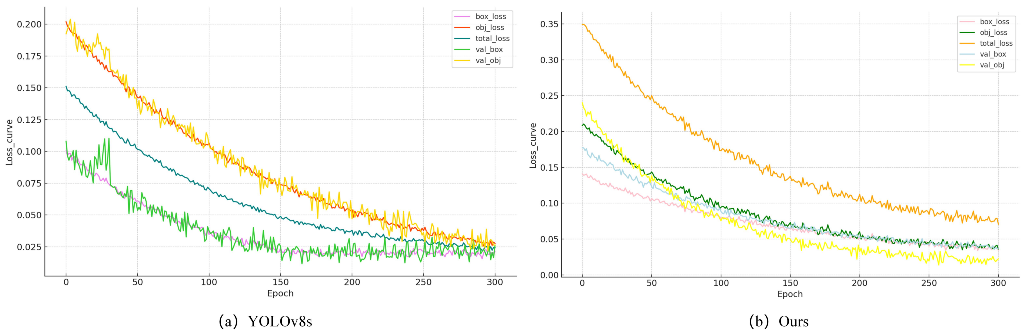

Figure 11. Loss curve of (**a**) YOLOv8s and (**b**) Ours.

This consistent convergence behavior observed across both training and validation datasets provides fundamental evidence of the statistical robustness of the proposed framework, indicating that the learning process is stable and reproducible. This behavior indicates effective learning of 6000 discriminative features for MO defects. Notably, the proposed model exhibits smoother loss curves with reduced oscillation compared to YOLOv8-s, which reflects enhanced numerical stability during optimization. Furthermore, the minimal gap maintained between the training and validation loss trajectories demonstrates superior generalization capability and statistical robustness, ensuring consistent detection performance even under potential data noise or unseen defect variations.

Grad-CAM (Gradient-weighted Class Activation Mapping) was utilized to analyze the activation patterns of specific layers in both the proposed improved model and the baseline model, offering qualitative insights into the image regions each model prioritizes during classification [44]. By highlighting the areas of an input image that are most influential for the final prediction, visualized as heatmaps, Grad-CAM helps interpret the models' decision-making processes.

It should be noted that the heatmaps shown in the top and bottom rows of Figure 12 correspond to the same defect instance, enabling a direct and fair comparison of attention distributions between the proposed method and the baseline model. As shown in the top row, the resulting heatmaps exhibit strong activations that are highly concentrated within the defective regions, particularly along defect-cluster boundaries and fine crack structures. This focused activation pattern indicates that the proposed attention-enhanced model effectively emphasizes defect-related features while suppressing background magnetic interference.

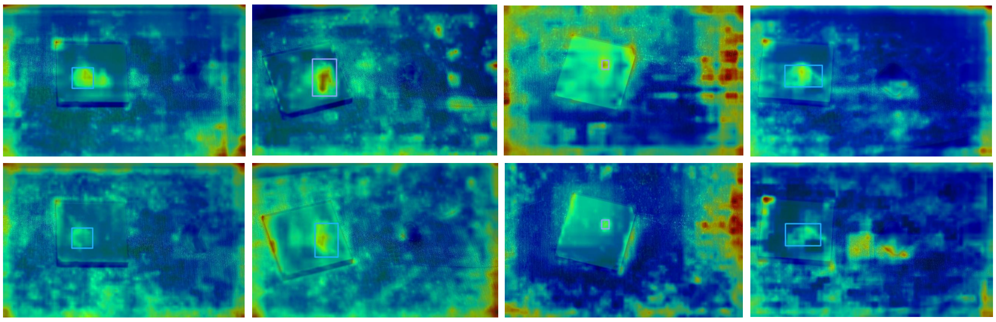

Figure 12. Comparison of MO defect heatmaps across models.

In contrast, the bottom row presents comparatively weaker and more dispersed activation responses, with limited emphasis on the actual defect areas. Owing to the absence of adaptive attention mechanisms, the baseline model tends to rely on more diffuse or background-associated features, resulting in less precise localization of defect-specific cues. Collectively, these visual comparisons demonstrate that the proposed method produces a more discriminative and defect-focused attention distribution than the baseline, thereby providing qualitative evidence of its superior defect detection capability.

4. Discussion

The results presented in the preceding section demonstrate the robust performance of the proposed model across a wide spectrum of industrial defect scenarios. The framework accurately detects and localizes defects of varying sizes, densities, and complexities—including overlapping or occluded targets—exhibiting strong generalization capabilities and high suitability for practical industrial deployment.

Nevertheless, several limitations were observed. Detection accuracy slightly declines for densely clustered micro-defects, particularly in the crack and scratch categories, which share high morphological similarity and are therefore prone to mutual misclassification. Patch-type defects also remain challenging due to low contrast with background textures and sensitivity to variations in illumination and color. These observations suggest that incorporating multi-scale feature enhancement or context-aware attention mechanisms could further improve detection reliability for these complex defect types.

Analysis of the training and validation loss trajectories confirms that the proposed model achieves superior stability and convergence relative to the baseline YOLOv8 model. The smoother loss curves indicate enhanced generalization and effective extraction of discriminative features from magneto-optical images. These results underscore the critical role of the customized network architecture and integrated attention mechanisms in strengthening model robustness.

In conclusion, the proposed approach effectively balances detection accuracy, computational efficiency, and deployment feasibility. Future research should prioritize addressing the identified limitations, with particular emphasis on improving the detection of small, densely clustered, and visually ambiguous defects.

5. Conclusions

This study presents a real-time deep learning framework for steel plate defect detection that integrates alternating magnetic excitation with MOI. By leveraging physically informed magnetic features beyond conventional visual cues, the network effectively identifies subtle and low-contrast defects. Key innovations include the integration of S-GhostConv for efficient feature extraction and an adaptive multi-scale feature fusion mechanism, which together enhance the representation of small defects and overall robustness. Experimental results demonstrate that the proposed model achieves an mAP@0.5 of 0.990 with a real-time inference speed of 130 FPS, outperforming the YOLOv8-s baseline by approximately 7% in detection accuracy under comparable computational conditions. Ablation studies further confirm the contribution of each module and validate the efficiency of the lightweight real-time architecture. Taken together, these results establish the framework as a high-performance solution for industrial defect inspection. Future work will focus on improving generalization and extending applicability to a wider range of operational conditions.

Author Contributions: Conceptualization, Q.N. and K.X.; methodology, C.Z.; software, Y.C.; validation, C.Z., D.C. and Q.N.; formal analysis, Q.N. and C.Z.; investigation, C.Z. and Q.N.; resources, Y.T. and K.X.; data curation, D.C. and C.Z.; writing—original draft preparation, C.Z.; writing—review and editing, Y.T. and K.X.; visualization, Q.N. and C.Z.; supervision, L.L.; project administration, K.X. and Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (62303094), the Program for Science & Technology Innovation Talents in Universities of Henan Province (25HASTIT009), the Key Research and Development Program of Henan Province of China (251111221700), the Key Scientific and Technological Project of Henan Province of China (252102221049).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are not available due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MOI	Magneto-optical Imaging
NDT	Non-destructive Testing
cGAN	Conditional Generative Adversarial Network
PANet	Path Aggregation Network
mAP	Mean Average Precision
CAM	Channel Attention Mechanism
SAM	Spatial Attention Mechanism
FPS	Frames Per Second
FLOPs	Floating point operations
YOLO	You Only Look Once
ROI	Region of Interest
CMOS	Complementary Metal-Oxide-Semiconductor
LED	Light Emitting Diode
CNN	Convolutional Neural Network
PCA	Principal component analysis
BP	Backpropagation neural network
SVM	Support vector machine
SGD	Stochastic gradient descent
SPPF	Spatial Pyramid Pooling - Fast
BN	Batch Normalization
ReLU	Rectified Linear Unit
MSE	Mean Squared Error
LPIPS	Learned Perceptual Image Patch Similarity
AC	Alternating Current
Grad-CAM	Gradient-weighted Class Activation Mapping
FEA	Finite Element Analysis

References

1. Li, J.; Zhan, D.; Jiang, Z.; Zhang, H.; Yang, Y.; Zhang, Y. Progress on improving strength-toughness of ultra-high strength martensitic steels for aerospace applications: A review. *J. Mater. Res. Technol.* **2023**, *23*, 172–190. [CrossRef]
2. Segovia Ramírez, I.; García Márquez, F.P.; Papaelias, M. Review on additive manufacturing and non-destructive testing. *J. Manuf. Syst.* **2023**, *66*, 260–286. [CrossRef]
3. Bajaj, P.; Hariharan, A.; Kini, A.; Kürnsteiner, P.; Raabe, D.; Jägle, E.A. Steels in additive manufacturing: A review of their microstructure and properties. *Mater. Sci. Eng. A* **2020**, *772*, 138633. [CrossRef]
4. Lou, W.; Shen, C.; Zhu, Z.; Liu, Z.; Shentu, F.; Xu, W.; Lang, T.; Zhang, Y.; Jing, Z.; Peng, W. Internal Defect Detection in Ferromagnetic Material Equipment Based on Low-Frequency Electromagnetic Technique in 20# Steel Plate. *IEEE Sens. J.* **2018**, *18*, 6540–6546.
5. Gupta, M.; Khan, M.A.; Butola, R.; Singari, R.M. Advances in applications of Non-Destructive Testing (NDT): A review. *Adv. Mater. Process. Technol.* **2022**, *8*, 2286–2307. [CrossRef]
6. Gao, X.; Zhen, R.; Xiao, Z.; Katayama, S. Modeling for detecting micro-gap weld based on magneto-optical imaging. *J. Manuf. Syst.* **2015**, *37*, 193–200. [CrossRef]
7. Zhang, Z.; Wen, G.; Chen, S. Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding. *J. Manuf. Process.* **2019**, *45*, 208–216. [CrossRef]
8. He, G.; Gao, X.; Yang, H. AETMC-FCVT: An end-to-end welding defect detection and classification method based on magneto-optical infrared bi-imaging system. *Mech. Syst. Signal Process.* **2025**, *224*, 112058. [CrossRef]
9. Shu, Z.; Wu, A.; Si, Y.; Dong, H.; Wang, D.; Li, Y. Automated identification of steel weld defects, a convolutional neural network improved machine learning approach. *Front. Struct. Civ. Eng.* **2024**, *18*, 294–308. [CrossRef]
10. Machado, M.A. Eddy Currents Probe Design for NDT Applications: A Review. *Sensors* **2024**, *24*, 5819. [CrossRef]
11. Xu, K.; Chai, D.; Shi, J.; Li, P.; Ai, J.; Wang, L. Eddy Current Magneto-Optical Imaging Based on Phase Change for CFRP Defects. *IEEE Sens. J.* **2025**, *25*, 13850–13858. [CrossRef]
12. Gao, X.; Zhang, Q.; Wang, C. Magneto-optical Imaging Detection and Strong Classification of Weld Defects in Rotating Magnetic Field. *J. Mech. Eng.* **2019**, *55*, 61.
13. Xie, WN.; Sun, X.Y.; Ma, W.F. A Light Weight Multi-Scale Feature Fusion Steel Surface Defect Detection Model Based on YOLOv8. *Meas. Sci. Technol.* **2024**, *35*, 055017. [CrossRef]
14. Lu, M.; Sheng, W.Q.; Zou, Y.; Chen, Y.T.; Chen, Z.G. WSS-YOLO: An Improved Industrial Defect Detection Network for Steel Surface Defects. *Measurement* **2024**, *236*, 115060. [CrossRef]
15. Gao, X.; Zhou, X.; Wang, C.; Ma, N.; Zhang, Y.; You, D. Skin depth and detection ability of magneto-optical imaging for weld defects in alternating magnetic field. *J. Manuf. Syst.* **2020**, *55*, 44–55. [CrossRef]
16. He, X.; Wang, T.; Wu, K.; Liu, H. Automatic defects detection and classification of low carbon steel WAAM products using improved remanence/magneto-optical imaging and cost-sensitive convolutional neural network. *Measurement* **2021**, *173*, 108633. [CrossRef]
17. Liu, Q.; Ye, G.; Gao, X.; Zhang, Y.; Gao, P.P. Magneto-optical imaging nondestructive testing of welding defects based on image fusion. *NDT E Int.* **2023**, *138*, 102887. [CrossRef]
18. Gao, X.; Du, L.; Ma, N.; Zhou, X.; Wang, C.; Gao, P.P. Magneto-optical imaging characteristics of weld defects under alternating and rotating magnetic field excitation. *Opt. Laser Technol.* **2019**, *112*, 188–197. [CrossRef]
19. Li, Z.; Xiao, L.; Shen, M.; Tang, X. A lightweight YOLOv8-based model with Squeeze-and-Excitation Version 2 for crack detection of pipelines. *Appl. Soft Comput.* **2025**, *177*, 113260. [CrossRef]
20. Qiao, T.; Chen, Y.; Zhou, X.; Shi, R.; Shao, H.; Shen, K.; Luo, X. CSC-Net: Cross-Color Spatial Co-Occurrence Matrix Network for Detecting Synthesized Fake Images *IEEE Trans. Cogn. Dev. Syst.* **2024**, *16*, 369–379. [CrossRef]
21. Li, Y.; Gao, X.; Zhang, Y. Detection and Strong Classification of Natural Weld Defects by Magneto-Optical Imaging Under Rotating Magnetic Field Excitation. *IEEE Access* **2024**, *12*, 161805–161819. [CrossRef]
22. Hayati, R.; Munawar, A.A.; Lukitaningsih, E.; Earlia, N.; Karma, T.; Idroes, R. Combination of PCA with LDA and SVM classifiers: A model for determining the geographical origin of coconut. *Case Stud. Chem. Environ. Eng.* **2024**, *9*, 100552. [CrossRef]
23. Chu, Y.; Cao, J.; Ding, W.; Huang, J.; Ju, H.; Cao, H.; Liu, G. Hyperspectral image classification using feature fusion fuzzy graph broad network. *Inf. Sci.* **2025**, *689*, 121504. [CrossRef]
24. Chen, F.; Deng, M.; Gao, H.; Yang, X.; Zhang, D. AP-Net: A metallic surface defect detection approach with lightweight adaptive attention and enhanced feature pyramid. *Clust. Comput.* **2024**, *27*, 3837–3851. [CrossRef]
25. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In *Computer Vision—ECCV 2018*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 3–19.
26. Ma, N.; Gao, X.; Tian, M.; Wang, C.; Zhang, Y.; Gao, P.P. Magneto-Optical Imaging of Arbitrarily Distributed Defects in Welds under Combined Magnetic Field. *Metals* **2022**, *12*, 1055. [CrossRef]

27. Wang, L.; Yi, S.; Yu, Y.; Gao, C.; Samali, B. Automated ultrasonic-based diagnosis of concrete compressive damage amidst temperature variations utilizing deep learning. *Mech. Syst. Signal Process.* **2024**, *221*, 111719. [CrossRef]
28. Wang, X.-B.; Zhang, X.; Li, Z.; Wu, J. Ensemble extreme learning machines for compound-fault diagnosis of rotating machinery. *Knowl.-Based Syst.* **2020**, *188*, 105012. [CrossRef]
29. Wang, Z.; Zhang, S.; Chen, Y.; Xia, Y.; Wang, H.; Jin, R.; Wang, C.; Fan, Z.; Wang, Y.; Wang, B. Detection of small foreign objects in Pu-erh sun-dried green tea: An enhanced YOLOv8 neural network model based on deep learning. *Food Control* **2025**, *168*, 110890. [CrossRef]
30. He, J.; Xu, G.; Wang, J.; Yan, F.; Ge, J.; Xiao, Z. Rail crack size measurement based on magneto-optical imaging in multi-layer lift-off. *NDT E Int.* **2025**, *152*, 103323. [CrossRef]
31. Zhao, X.; Yang, T.; Li, B.; Zhang, X. SwinGAN: A dual-domain Swin Transformer-based generative adversarial network for MRI reconstruction. *Comput. Biol. Med.* **2023**, *153*, 106513. [CrossRef]
32. Huang, X.; Zhu, J.; Huo, Y. SSA-YOLO: An Improved YOLO for Hot-Rolled Strip Steel Surface Defect Detection. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 5040017. [CrossRef]
33. Jin, J.; Qin, Z.; Yu, D.; Li, Y.; Liang, J.; Chen, C.L.P. Regularized discriminative broad learning system for image classification. *Knowl.-Based Syst.* **2022**, *251*, 109306. [CrossRef]
34. Lv, P.; Zhang, Q.; Wu, M.; Deveci, M.; Shi, L.; Sun, Y.; Zhong, K.; Wang, J. Intelligent extraction of medical entity relationship based on graph neural network and optimization strategy. *Knowl.-Based Syst.* **2024**, *294*, 111735. [CrossRef]
35. Yang, T.; Bai, X.; Cui, X.; Gong, Y.; Li, L. GraformerDIR: Graph convolution transformer for deformable image registration. *Comput. Biol. Med.* **2022**, *147*, 105799. [CrossRef] [PubMed]
36. Ding, P.; Zhan, H.; Yu, J.; Wang, R. A bearing surface defect detection method based on multi-attention mechanism YOLOv8. *Meas. Sci. Technol.* **2024**, *35*, 086003. [CrossRef]
37. Li, Y.; Jin, J.; Geng, Y.; Xiao, Y.; Liang, J.; Chen, C.L.P. Discriminative elastic-net broad learning systems for visual classification. *Appl. Soft Comput.* **2024**, *155*, 111445. [CrossRef]
38. Cao, H.; Cao, J.; Chu, Y.; Wang, Y.; Liu, G.; Li, P. Global-local manifold embedding broad graph convolutional network for hyperspectral image classification. *Neurocomputing* **2024**, *602*, 128271. [CrossRef]
39. Tao, Y.; Karimian, H.; Shi, J.; Wang, H.; Yang, X.; Xu, Y.; Yang, Y. MobileYOLO-Cyano: An enhanced deep learning approach for precise classification of cyanobacterial genera in water quality monitoring. *Water Res.* **2025**, *285*, 124081. [CrossRef]
40. Li, W.; Yu, Y.; Meng, F.; Duan, J.; Zhang, X. An image fusion and U-Net approach to improving crop planting structure multi-category classification in irrigated area. *Int. J. Food Sci.* **2023**, *45*, 185–198. [CrossRef]
41. Guo, X.; Li, W.-J.; Qiao, J.-F. A modular neural network with empirical mode decomposition and multi-view learning for time series prediction. *SSRN Electron. J.* **2023**. [CrossRef]
42. Han, K.; Wang, Y.; Tian, Q.; Guo, J.; Xu, C.; Xu, C. GhostNet: More Features From Cheap Operations. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Piscataway, NJ, USA, 2020; pp. 1577–1586.
43. Liao, Y.; Li, L.; Xiao, H.; Xu, F.; Shan, B.; Yin, H. YOLO-MECD: Citrus Detection Algorithm Based on YOLOv11. *Agronomy* **2025**, *15*, 687. [CrossRef]
44. Chen, T.; Sun, L.; Wang, Z.; Wang, Y.; Zhao, Z.; Zhao, P. An enhanced nonlinear iterative localization algorithm for DV_Hop with uniform calculation criterion. *Ad Hoc Netw.* **2021**, *111*, 102327. [CrossRef]