

Article

PDA-YOLO: An Early Detection Method for Egg Fertilization Rate Based on Position-Decoupled Attention

Yifan Zhou ^{1,2,3}, Zhengxiang Shi ³ , Geqi Yan ⁴, Haiqing Peng ^{1,2}, Fuwei Li ^{1,2}, Wei Liu ^{1,2} and Dapeng Li ^{1,2,*}

¹ Poultry Institute, Shandong Academy of Agricultural Sciences, Jinan 250100, China; mondayzyf@cau.edu.cn (Y.Z.); 15562717588@163.com (H.P.); lifuwei1224@163.com (F.L.); lwteam@126.com (W.L.)

² Shandong Provincial Key Laboratory of Livestock and Poultry Breeding, Jinan 250100, China

³ College of Water Resources and Civil Engineering, China Agricultural University, Beijing 100083, China; shizhx@cau.edu.cn

⁴ College of Animal Science and Technology, Shandong Agricultural University, Taian 271018, China; yangeqi@sdaa.edu.cn

* Correspondence: lidp1990@126.com

Abstract

This study addresses the inefficiencies, subjectivity, and poor adaptability to lighting variations inherent in traditional candling methods used in large-scale egg incubation. We developed a high-throughput transmissive imaging system capable of capturing 30 eggs simultaneously. Based on this system, we propose PDA-YOLO, an enhanced YOLOv8-based object detection model featuring a position-decoupled attention strategy. Specifically, a lightweight C2f-SE module is integrated into the backbone to amplify subtle feature responses in low-contrast regions, while a CBAM is deployed prior to the detection head to mitigate background clutter through precise spatial attention. Experimental results on a self-constructed Hailan White egg dataset show that at the critical 60 h incubation stage, PDA-YOLO achieves a Recall of 91.5% and an $mAP@0.5$ of 97.4%, outperforming the YOLOv8 baseline while maintaining a real-time inference speed of 62.1 FPS. Grad-CAM visualizations confirm the model's ability to focus on vascular textures and suppress noise. Furthermore, the model demonstrates robust performance under varying illumination (180–540 lumens), effectively mitigating missed detections in low light and recognition degradation from over-exposure. This work provides a scalable, real-time solution for non-destructive, early-stage detection of poultry health and fertilization status in commercial hatcheries.

Keywords: fertilized egg incubation; fertilization rate detection; position-decoupled attention; robustness; YOLOv8



Academic Editor: Kaiying Wang

Received: 12 February 2026

Revised: 23 March 2026

Accepted: 31 March 2026

Published: 2 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the growing global demand for high-quality protein, the economic significance of the poultry industry has become increasingly prominent. As a core link in the production chain, the efficiency of fertilized egg incubation directly determines the economic returns of poultry enterprises. The incubation process, which involves precise control of temperature, humidity, and ventilation, supports normal embryonic development [1,2]. However, due to factors such as egg storage conditions (e.g., temperature, humidity, and storage duration) and the health status of breeder flocks, a proportion of unfertilized eggs or early dead embryos inevitably exist among eggs set for incubation [3,4]. If these non-viable eggs are not removed in a timely manner, they not only occupy valuable incubation space and

consume energy resources, but may also deteriorate and burst under high-temperature, high-humidity conditions, contaminating surrounding healthy eggs and consequently reducing the overall hatchability [5]. Therefore, the rapid and accurate identification and removal of fertilized eggs at the early stage of incubation are of great significance for optimizing production management and reducing costs.

Timely identification and removal of unfertilized eggs before and during the early incubation stages have long been a key issue in poultry hatchery production that remains unresolved [6]. Although the chicken embryo development process has been widely studied as a classical biological model [7], in actual industrial incubation conditions, fertilization status and embryo viability information are primarily available from the 5th to the 7th day of incubation [8]. Currently, the incubation industry predominantly relies on traditional manual candling techniques, which utilize strong transmitted light to observe the internal characteristics of eggs. While this method is simple to operate, it suffers from significant drawbacks: its accuracy is highly dependent on the operator's experience, and subjective judgment often leads to a high rate of misclassification; furthermore, it is labor-intensive, inefficient, and prolonged exposure to intense light may have adverse effects on the inspectors' vision health [9,10]. As a result, the development of automated, non-invasive intelligent detection equipment has become an urgent need in the industry.

To overcome the limitations of manual inspection, various non-destructive detection techniques have been explored. Among them, spectroscopic techniques have demonstrated high early-stage detection accuracy [11–15]. For instance, Schreuder et al. achieved 93–95% classification accuracy on the 7th day of incubation using near-infrared spectroscopy [16]; Ghaderi et al. and Adegbenjo et al. successfully identified fertilized eggs on the 0th and 5th days of incubation, respectively, with accuracies exceeding 93% [17,18]; moreover, Wang et al. achieved over 99% detection accuracy on the 3rd day of incubation by monitoring respiratory and growth indicators of fertilized eggs [19]. However, despite the ability of spectroscopic and dielectric techniques to perform early-stage detection, their high equipment costs, complex systems, slow detection speeds, and limited throughput present significant challenges for large-scale industrial deployment.

In contrast, computer vision-based detection technologies, with their low cost, non-contact operation, and high throughput, exhibit great potential for industrial application [20–23]. Li et al. designed an image acquisition system and optimized a YOLOv7 algorithm, achieving 98.10% detection accuracy for duck eggs [10]; Geng et al. combined transfer learning and neural networks, achieving 99.5% accuracy on the 5th day of incubation [24]; Çevik et al. employed Mask R-CNN, achieving 100% detection accuracy on the 3rd day of incubation [25]. These studies demonstrate the excellent performance of deep learning in feature extraction and pattern recognition. Despite the promising results achieved by the aforementioned deep learning methods, several critical challenges persist in their transition to industrial-scale applications. First, environmental adaptability remains limited. Most existing studies focus on single-egg detection under idealized lighting, whereas standard models often lack the robustness to handle the illumination fluctuations—such as local overexposure or light attenuation—typically encountered in hatcheries. Second, extracting subtle biological features is inherently difficult. In the early stages of incubation (e.g., within 72 h), embryonic vascular patterns are extremely faint, making it arduous for conventional Convolutional Neural Networks (CNNs) to distinguish these vital features from complex background noise. Finally, there is a distinct trade-off between detection efficiency and accuracy. While the integration of complex attention mechanisms can enhance precision, it significantly increases computational overhead, failing to meet the real-time inference requirements for high-throughput batch inspection in industrial pipelines.

To address these issues, this study proposes an integrated solution combining low-cost acquisition hardware with interference-resistant detection algorithms. The primary contributions of this work are as follows: First, a high-throughput image acquisition system was developed. Equipped with a customized LED light source and a darkroom structure, the system enables the simultaneous capture of images from 30 Hailan White fertilized eggs, providing a hardware foundation for industrial-scale batch inspection. Second, an improved object detection model (Model D) based on YOLOv8 is proposed. By incorporating a lightweight Squeeze-and-Excitation (SE) module into the backbone network and integrating a Convolutional Block Attention Module (CBAM) into the detection head, the model enhances the perception of subtle embryonic features while suppressing interference from complex illumination fluctuations. Third, a systematic evaluation of model robustness was conducted. The stability of the model was verified under varying illumination intensities (180, 360, and 540 lumens), demonstrating its adaptability to uneven lighting environments commonly encountered in real-world hatcheries. Finally, an optimal balance between accuracy and efficiency was achieved at the critical 60 h incubation stage. These findings provide technical support and a theoretical foundation for the development of intelligent equipment for early-stage embryonic detection.

Finally, an optimal balance between accuracy and efficiency was achieved at the critical 60 h incubation stage, providing a scalable, real-time industrial solution that offers technical support for the development of intelligent equipment for early-stage embryonic detection.

2. Materials and Methods

2.1. Overall Architecture of the Proposed Model

To tackle the challenges in fertilized egg incubation detection, such as substantial illumination variations, weak target texture features, and stringent real-time performance requirements, this study presents an enhanced YOLOv8-based object detection method incorporating a position-decoupled attention mechanism. The term “position-decoupled” refers to the separation of spatial position information from the attention mechanism, allowing the model to focus on the most relevant features without being influenced by the exact position of the target. The proposed approach aims to enhance the detection accuracy and robustness for weak-texture targets under complex illumination conditions while maintaining real-time inference performance.

This study is based on the YOLOv8 object detection framework [26], with its network architecture specifically optimized for the task. As an efficient single-stage object detection algorithm, YOLOv8 has been widely applied in real-time application scenarios, including agricultural vision detection [27]. While YOLOv8 demonstrates stable performance under standard conditions, it faces challenges in controlled industrial environments, such as fertilized egg incubation, where illumination intensity varies significantly and target appearance features are relatively homogeneous. These challenges can negatively impact detection accuracy, especially in fluctuating lighting conditions or when the target lacks distinctive texture features, which are common in such environments.

For the common challenges of high target density and complex backgrounds in batch fertilized egg detection, related studies have shown that in similar unstructured and dense scenarios, traditional algorithms often struggle to achieve a balance between detection accuracy and inference speed. Introducing attention mechanisms has been proven to be an effective means of enhancing feature extraction capability [28]. Meanwhile, to cope with potential illumination non-uniformity and target scale variations in incubation detection environments, detection models are required to comprehensively capture both channel-wise and spatial information in feature maps to ensure robustness [29].

Specifically, this study proposes a position-decoupled attention deployment strategy as follows:

- (1) Internal backbone enhancement: The C2f modules within the backbone are replaced with a lightweight channel attention mechanism (C2f-SE) to enhance channel-wise feature representation across multiple scales;
- (2) Backbone terminal enhancement: A spatially aware attention mechanism (CBAM) is introduced at the deepest output layer of the backbone (before the SPPF module) to perform spatial calibration of high-level semantic features prior to feature fusion.

The neck structure is kept unchanged. Through the above design, the proposed method effectively improves detection performance under complex illumination conditions without significantly increasing model parameters or computational overhead. The overall architecture of the proposed model is illustrated in Figure 1.

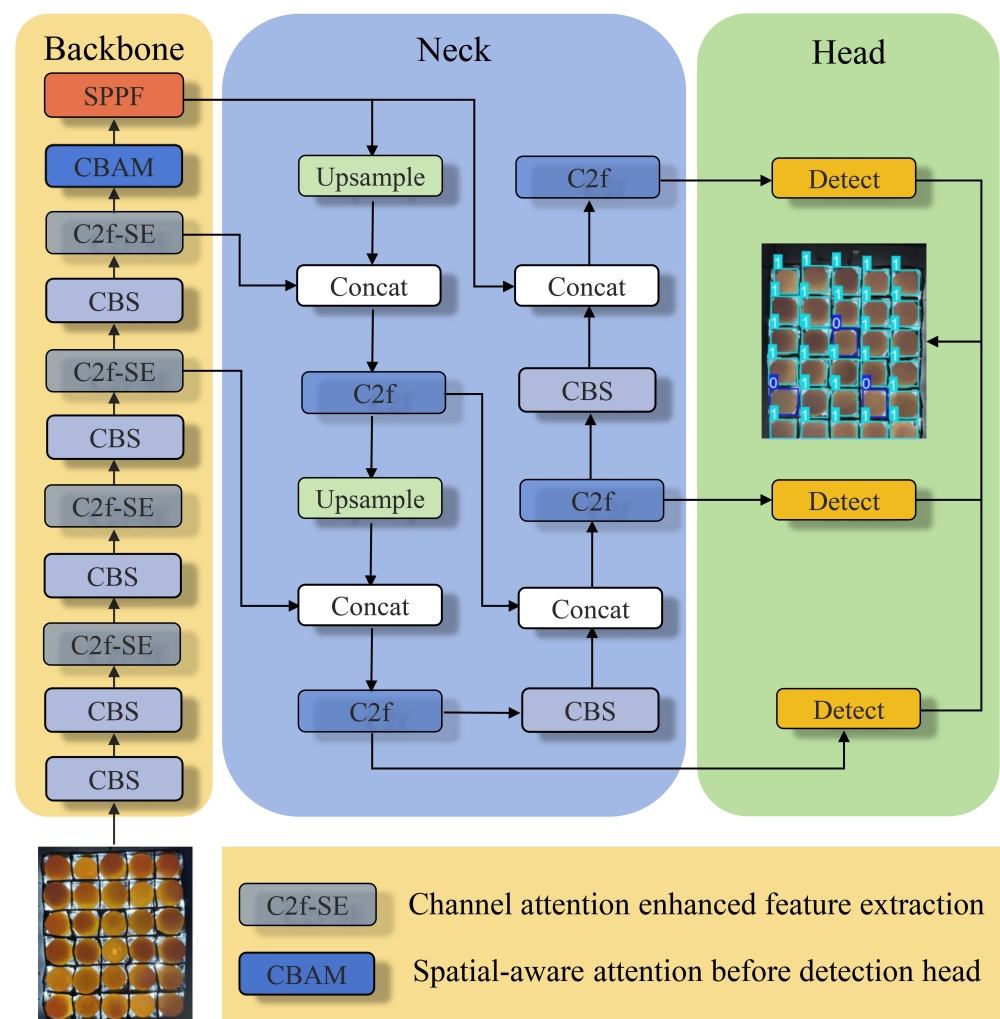


Figure 1. Overall architecture of the improved YOLOv8 model based on a position-decoupled attention mechanism.

2.2. C2f-SE Module

2.2.1. Task-Oriented Motivation for SE

In fertilized egg incubation detection tasks, the overall appearance of targets is highly similar, and the discriminative information among different categories is mainly reflected in subtle differences in color distribution and local texture. Under complex illumination

conditions, such fine-grained differences are easily affected by factors such as shadows, reflections, and uneven brightness, leading to unstable feature response distributions.

During the feature extraction process of convolutional neural networks, different feature channels often contribute unequally to target discrimination. However, conventional feature extraction structures typically assign equal weights to all channels, lacking explicit modeling of channel importance. In weak-texture and low-contrast scenarios, this uniform treatment tends to obscure feature channels that are more critical for target discrimination, thereby degrading overall detection performance. For tasks requiring the classification of fine-grained features, introducing attention mechanisms to enhance inter-channel dependencies can effectively improve the model's perception of subtle features [30].

Therefore, from the perspective of task characteristics, incorporating an attention mechanism capable of adaptively recalibrating channel features helps strengthen the representation of key semantic features and improve the robustness of the model under complex illumination conditions.

2.2.2. Structure of the C2f-SE Module

The Squeeze-and-Excitation (SE) attention mechanism explicitly models the inter-dependencies among feature channels and adaptively recalibrates channel-wise feature responses, thereby enhancing the model's discriminative capability for key features [31]. Given the limitation of the original C2f module in YOLOv8, which treats all channel features equally, this study adopts recent improvement strategies by embedding the SE module into the C2f structure to construct the C2f-SE module [28].

In terms of structural implementation, as illustrated in the figure, the improved C2f-SE module first processes the input features through convolution and splits them into two branches. One branch enters a deep feature extraction path composed of N stacked Bottleneck modules, while the other branch is directly forwarded to the concatenation layer. After the two branches are concatenated at the Concat layer, a second convolution layer is applied to perform preliminary feature fusion. Instead of directly outputting the fused features, the module feeds them into an SE attention branch for feature recalibration. Specifically, the feature maps are first compressed into a global feature vector through Global Average Pooling, which aggregates spatial information. Subsequently, the channel dimension is reduced to $C/16$, a ratio selected based on design principles from Hu et al., CVPR 2018, which demonstrated that this reduction ratio offers an optimal balance between accuracy and computational complexity. This reduction is achieved through a convolution layer with a ReLU activation function, followed by another convolution layer with a Sigmoid activation function to restore the channel dimension to C and generate normalized channel-wise weight coefficients. The reduction ratio of $C/16$ is chosen based on the design principles proposed by Hu et al. [31], which demonstrated that this ratio provides an optimal balance between accuracy and computational complexity.

Through this design, the C2f-SE module effectively enhances the channel representation capability of multi-scale features without altering the original feature extraction structure. A schematic illustration of the C2f-SE module is shown in Figure 2.

2.2.3. Rationale for Backbone Modification

During the backbone feature extraction stage, constructing high-quality semantic representations is more critical than precise spatial localization. Channel attention mechanisms focus on modeling the semantic correlations among different feature channels, enabling adaptive selection of feature layers that are beneficial for target discrimination. Integrating the SE module into the backbone not only enhances the responses of key semantic features but also effectively suppresses background noise, such as specular highlights and clutter

interference in candling environments, thereby improving the recognition and extraction of important features.

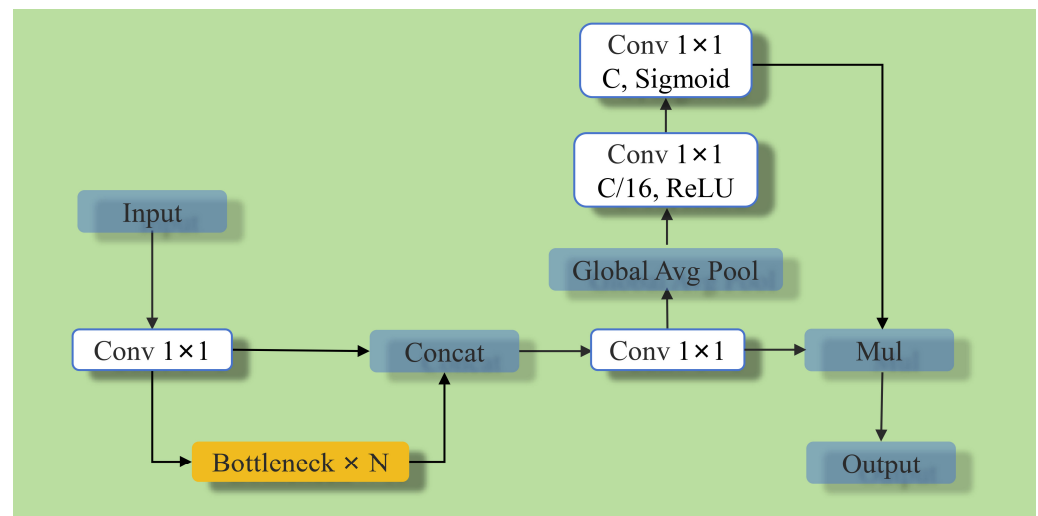


Figure 2. Schematic illustration of the C2f-SE module.

Furthermore, owing to the dimensionality reduction design within the module (reduction ratio = 16), the SE module introduces only a negligible increase in model parameters. Seamlessly embedding it into the C2f module significantly enhances feature extraction capability while maximally preserving the overall computational efficiency and real-time inference performance of the model, thereby achieving an excellent balance between detection accuracy and processing speed.

2.3. Neck Module

During the feature fusion stage, this study adopts the original multi-scale feature fusion structure of YOLOv8 to effectively integrate features at different scales output by the backbone. This structure is based on a bidirectional feature propagation mechanism combining top-down and bottom-up pathways, which introduces high-level semantic information while preserving spatial details, thereby enhancing the model's capability to detect objects of varying scales.

Specifically, the multi-scale features extracted from the backbone are first propagated through a top-down pathway to transmit semantic information, allowing high-level semantic features to be gradually fused into lower-level features. Subsequently, a bottom-up pathway is employed to further reinforce spatial detail information, achieving bidirectional complementarity in feature representation. Through this multi-scale feature fusion process, the model ultimately generates fused feature maps at multiple scales, which serve as multi-scale inputs for the subsequent detection head.

Given that this study focuses on the structural optimization and deployment strategy of attention mechanisms, no additional modules or parameter adjustments are introduced in the neck component. Keeping the neck structure unchanged helps ensure real-time performance while enabling a clear analysis of the impact of attention mechanisms in the backbone and detection head on overall detection performance.

2.4. CBAM

2.4.1. Task-Oriented Motivation for CBAM

In the deeper stages of convolutional neural networks, feature maps contain highly abstract semantic information; however, as network depth increases, the texture details

and spatial location information of fertilized egg targets tend to be weakened. In batch fertilized egg incubation detection scenarios, targets are typically densely distributed, and the differences between fertilized and unfertilized eggs are extremely subtle, such as faint vascular textures or shadow patterns. In addition, complex illumination conditions on practical production lines often lead to local overexposure or shadow occlusion, which severely interferes with the model's ability to focus on target regions.

Related studies have shown that, when dealing with agricultural detection tasks characterized by dense distributions of small targets, severe occlusion, and complex environments, the introduction of attention mechanisms can significantly enhance the extraction of critical features while effectively suppressing the influence of background noise [32]. Therefore, from the perspective of task requirements, introducing an attention mechanism that simultaneously performs channel-wise feature selection and spatial position calibration before deep features enter the feature fusion network is crucial for improving the robustness of fertilized egg detection.

2.4.2. CBAM and Its Deployment Strategy

The Convolutional Block Attention Module (CBAM) is a lightweight attention mechanism that performs progressive feature refinement and enhancement by jointly modeling channel attention and spatial attention on feature maps [33]. A schematic illustration of the CBAM is shown in Figure 3.

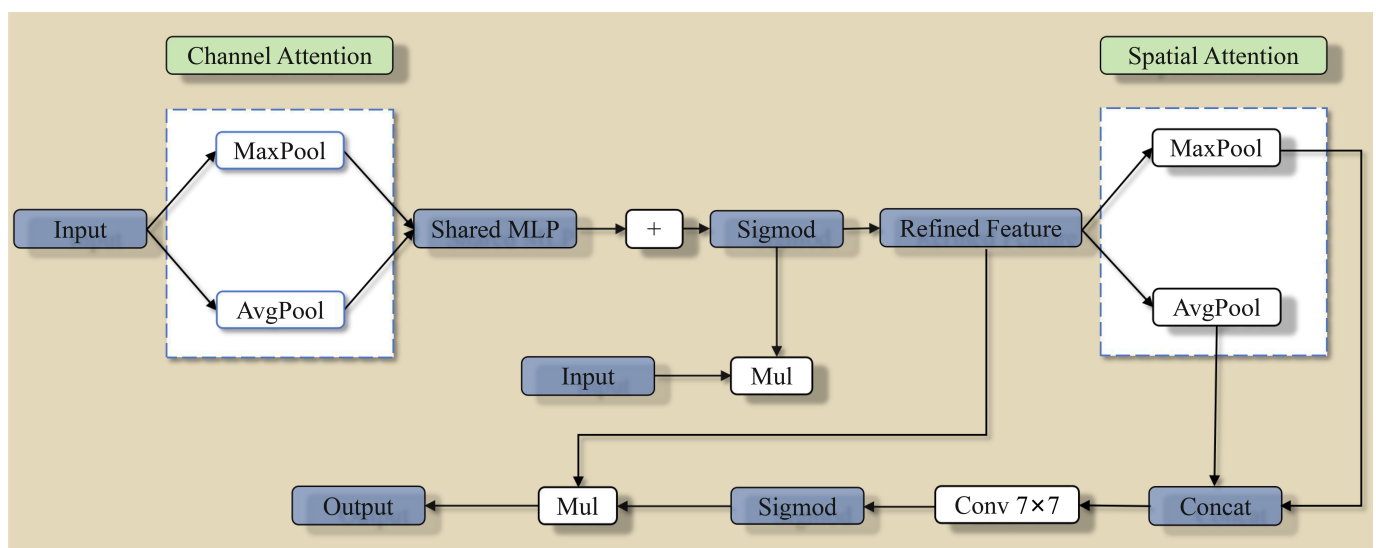


Figure 3. Schematic illustration of the CBAM.

Given the challenges of complex illumination conditions and densely distributed targets in fertilized egg detection, this study follows improvement strategies reported in related agricultural vision research and deploys the CBAM at the terminal stage of the backbone network, specifically between the last C2f-SE module and the SPPF module.

This position corresponds to the deepest feature level (P5) extracted by the backbone, where the feature maps have the lowest spatial resolution but contain the richest semantic information. At this stage, CBAM operates as follows: the spatial dimensions of the feature maps are first compressed through global average pooling and max pooling, and a shared multilayer perceptron (MLP) is employed to learn the relative importance of different channels. This process enables the model to identify the most discriminative “fertilization-related” feature channels during feature extraction, compensating for the limitations of standard convolution operations in modeling complex visual patterns [34].

Subsequently, based on the channel-refined features, spatial attention maps are generated through pooling operations along the channel dimension followed by convolution. These spatial attention maps guide the network to accurately focus on the effective regions of individual eggs within densely arranged batches, while suppressing interference from egg trays and specular highlights. As a result, spatial calibration of high-level semantic features is completed before entering the neck for feature fusion, thereby enhancing robustness under complex industrial illumination conditions [33].

2.4.3. Structure-Oriented Rationale

Before entering the detection head, the feature maps have already undergone multi-scale fusion and thus possess strong semantic discriminative capability. At this stage, the primary performance bottleneck of the model lies in the spatial localization accuracy of target regions. Deploying an attention mechanism with spatial awareness at this position helps further filter and enhance spatial responses before the final detection stage.

Compared with introducing complex spatial attention mechanisms within the backbone, placing CBAM before the detection head achieves significant spatial feature enhancement with relatively low computational overhead. This design improves detection stability under complex illumination conditions while preserving real-time inference performance. In combination with the channel attention-enhanced backbone design described earlier, this architecture forms an attention deployment strategy based on functional differences across network stages, enabling effective synergy between semantic feature enhancement and spatial localization accuracy improvement.

2.5. Dataset and Evaluation Metrics

2.5.1. Dataset and Experimental Environment

The experimental samples used in this study were collected in Jinan, Shandong Province, China, comprising a total of 1800 Hailan White chicken eggs, including 900 fertilized eggs and 900 unfertilized eggs. To ensure a sterile experimental environment and suppress microbial interference, the surface of each egg was uniformly disinfected by spraying a benzalkonium bromide solution with a concentration of 27–33 g/L. After disinfection, each egg was systematically labeled and placed in a constant temperature and humidity incubator (Dezhou Dingfeng Machinery Equipment Co., Ltd., Dezhou, China).

To optimize the predictive performance and robustness of the YOLOv8 model in practical production scenarios, a dataset with a balanced class distribution was constructed. The incubation environment was strictly controlled under standard conditions, with a temperature of 38 °C and a relative humidity of 60%. The incubator was equipped with 10 incubation trays, each divided into two groups, with 30 fertilized eggs placed in each group. Incubation was conducted in three separate batches. For each batch, incubation was carried out until the 7th day, after which the fertilization status of the eggs was confirmed through manual candling, and the results were used as the ground truth labels for the early-stage sample categories.

2.5.2. Image Acquisition System

The image acquisition system primarily comprised a computer, an industrial camera (a2A5320-23um, Basler Vision Technologies, Ahrensburg, Germany), a fully enclosed dark chamber, an LED array light source, and a customized egg tray. The imaging unit featured a Basler camera equipped with a CMOS sensor, capturing raw images at a resolution of 5320 × 4600 pixels. The camera was mounted vertically at a distance of approximately 500 mm directly above the egg tray, utilizing its high-resolution capabilities to capture subtle embryonic development features.

The dark chamber was constructed from 1 cm thick acrylic panels, with overall dimensions of 50 cm × 31 cm × 47 cm. An opaque polyester light-shielding curtain was installed at the conveyor opening to exclude ambient light interference, ensuring a controlled environment. Furthermore, the system utilized a black light-shielding tray arranged in a 5 × 6 layout, integrated with a trapezoidal support structure at the base and a rubber sponge pad on the top. This design physically stabilized the eggs while facilitating light convergence, effectively eliminating optical crosstalk between adjacent samples. The bottom-up LED array served as the transmissive light source, with its intensity adjustable across three levels (180, 360, and 540 lumens) to simulate varying industrial lighting conditions. All images were saved in JPG format and subsequently resized to 640 × 640 pixels during the preprocessing stage to meet the model input requirements. A schematic illustration of the system is presented in Figure 4.

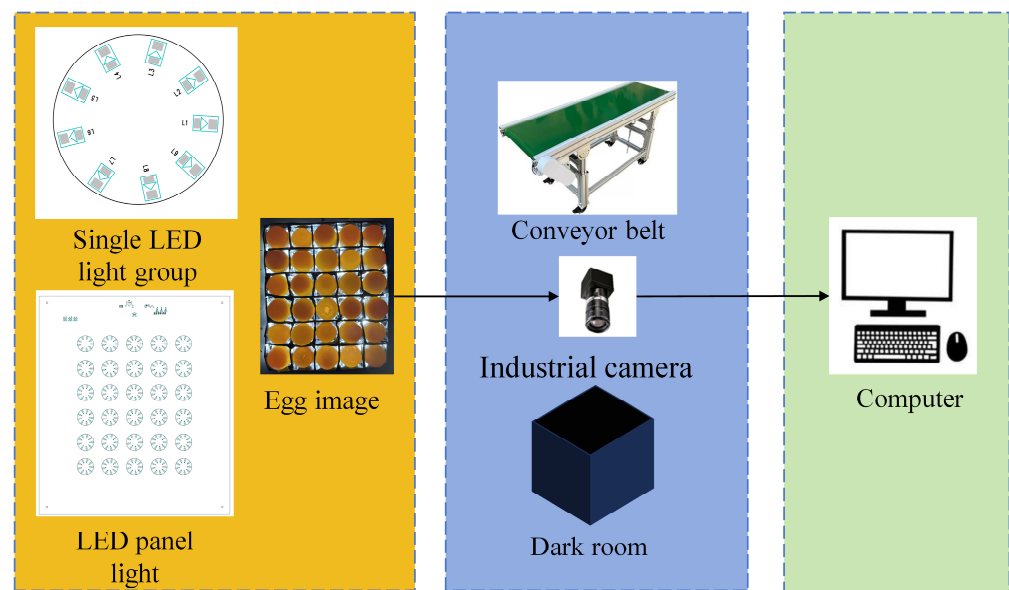


Figure 4. Schematic illustration of the image acquisition system.

2.5.3. Data Annotation and Preprocessing

Data annotation is a critical step for defining the locations of bounding boxes and category information, and its quality directly affects feature learning during model training as well as final performance evaluation [35]. In this study, the Labeling 1.8.6 software was used to manually annotate the 1800 selected egg images. All experimental images were acquired at 60 h of incubation. To ensure annotation accuracy, the fertilization status was confirmed via manual candling on the 7th day (168 h) and mapped back to the 60 h images as ground truth. The annotation process was conducted by an experienced researcher, followed by multiple rounds of manual review to ensure the consistency and reliability of the labels.

Regarding image preprocessing, all raw images were normalized and uniformly resized to 640 × 640 pixels to fit the model input. Notably, no offline or online data augmentation (such as Mosaic or geometric transformations) was employed during the training process. This decision was based on two considerations: First, the images were captured under a highly controlled environment with a consistent background, making the 1800 samples sufficiently representative for the task. Second, at the early 60 h incubation stage, embryonic physiological features are extremely subtle; standard augmentation techniques could introduce noise or distort these delicate image signals, potentially hindering the model's ability to extract authentic features.

To evaluate the effectiveness and generalization capability of the proposed model, the annotated dataset was randomly divided into training, validation, and test sets with a ratio of 70%:15%:15%. Specifically, the training set contained 1260 samples (630 fertilized eggs and 630 unfertilized eggs), while the validation and test sets each included 270 samples (135 fertilized eggs and 135 unfertilized eggs), ensuring balanced class distributions across all subsets.

2.5.4. Evaluation Metrics

To comprehensively evaluate the performance of the improved object detection model, this study considers detection accuracy, inference speed, and model complexity. Precision (P), Recall (R), mean Average Precision (mAP), number of parameters (Parameters), floating-point operations (FLOPs), and frames per second (FPS) are adopted as the core evaluation metrics. Among them, the calculation methods of P , R , and mAP are presented in Equations (1)–(3).

$$P = \frac{T_P}{T_P + F_P} \times 100\% \quad (1)$$

$$R = \frac{T_P}{T_P + F_P} \times 100\% \quad (2)$$

$$mAP = \frac{\sum_{i=1}^N \int_0^1 p(R) dR}{N} \times 100\% \quad (3)$$

2.5.5. Experimental Environment and Training Parameter Settings

To ensure the objectivity of algorithm evaluation, the experimental environment and training parameters were standardized in this study. Model training and optimization were performed using the AdamW optimizer (initial learning rate: 0.0005; batch size: 8) for a total of 300 epochs. The training process was supervised by a composite loss function, comprising Complete-IOU (CIoU) loss for bounding box regression, Binary Cross-Entropy (BCE) loss for classification, and Distribution Focal Loss (DFL) for boundary refinement. Crucially, consistent with the preprocessing strategy in Section 2.5.3, no additional online or offline data augmentation was applied. This strict adherence to the original data distribution ensures that the model extracts authentic physiological features under a standardized environment, thereby maintaining high fidelity for early embryonic detection. Detailed hardware and software configurations are presented in Table 1.

Table 1. Experimental environment and configuration.

Experimental Environment	Model/Parameter
CPU device	Intel(R) Core(TM) i9-12900H
GPU device	NVIDIA GeForce RTX 3060 (6 GB)
operating system	Windows 10
CUDA Version	CUDA 11.6/cuDNN 8.5
programming language	Python 3.8.19
Deep Learning Framework	Pytorch 1.10.0
IDE	PyCharm Professional Edition 2023.3

Furthermore, to ensure the statistical reliability of model evaluation and reduce the impact of randomness during training, each model variant was trained and evaluated through 10 independent runs using different random seeds. The corresponding performance metrics are reported in the format of mean $\pm$ standard deviation (SD). The calculations of

the mean and standard deviation were performed using IBM SPSS Statistics version 22.0 (IBM Corp., Armonk, NY, USA). To further verify whether the performance differences between models were statistically significant, a one-way ANOVA was conducted for the performance metrics (P , R , $mAP@0.5$, and $mAP@0.5:0.95$) of the six models.

3. Experimental Results and Analyses

3.1. Ablation Study

To ensure the statistical reliability and stability of the experimental results, each model variant was trained and evaluated through 10 independent runs. The performance metrics presented in Table 2 are reported as mean $\pm$ standard deviation. This approach accounts for the randomness in the training process, such as weight initialization and data shuffling, providing a robust basis for evaluating the proposed structural improvements.

Table 2. Results of the ablation experiment (mean $\pm$ standard deviation over 10 independent runs).

Model	$P/\%$	$R/\%$	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	Params (M)	FPS/(f·s ⁻¹)
A	93.01 $\pm$ 2.90	93.95 $\pm$ 1.99	95.21 $\pm$ 0.69	80.80 $\pm$ 0.78	5.78	18.3
B	89.95 $\pm$ 3.43	94.06 $\pm$ 2.01	97.21 $\pm$ 0.45	82.60 $\pm$ 0.86	3.02	24.8
C	92.77 $\pm$ 3.00	93.01 $\pm$ 2.32	96.93 $\pm$ 0.71	82.46 $\pm$ 0.64	3.01	82.6
D	91.45 $\pm$ 2.35	94.18 $\pm$ 1.77	97.07 $\pm$ 1.05	82.02 $\pm$ 0.99	3.03	62.1
E	89.93 $\pm$ 2.60	93.24 $\pm$ 0.95	96.17 $\pm$ 0.91	81.24 $\pm$ 1.27	3.04	26.2
F	91.11 $\pm$ 2.26	93.17 $\pm$ 2.26	95.80 $\pm$ 1.45	81.74 $\pm$ 1.21	3.03	21.3

Before conducting the ablation experiments, the detection performance of the baseline model (the original YOLOv8n) was first evaluated. The baseline model was trained and tested on the acquired dataset without introducing any structural modifications. It achieved a precision of 84.1%, a recall of 95.4%, an $mAP@0.5$ of 96.6%, and an $mAP@0.5:0.95$ of 82.4%. The false positive rate (FPR) was calculated to be 0.74%, indicating that while the model performs well overall, it still suffers from a minimal level of false detections under complex transmissive backgrounds and local reflection interference. These results highlight that relying solely on the original feature extraction structure is insufficient to effectively suppress background noise while maintaining a high recall rate. This provides a baseline for the subsequent introduction of attention mechanisms.

To verify the effectiveness of the proposed position-decoupled attention (PDA) mechanism, a systematic ablation study was conducted. Model A (EMA added to the baseline), Model B (CBAM added to the baseline), and Model C (SE added to the baseline) represent experimental groups where EMA, CBAM, and SE modules were individually introduced into the baseline. Model D (the proposed PDA-YOLO) integrates SE into the backbone and CBAM into the detection head. Model E integrates EMA into the backbone and SE into the detection head, while Model F integrates EMA into the backbone and CBAM into the detection head. Models E and F serve as comparative variants of the decoupled strategy to evaluate the trade-offs between precision and speed.

The single-module ablation experiments indicate that a single attention mechanism struggles to simultaneously balance detection accuracy and inference speed. Although Model B improved the average $mAP@0.5$ to 97.21%, its inference speed dropped significantly to 24.8 FPS, which fails to meet industrial real-time requirements (>60 FPS). Model C achieved a high throughput of 82.6 FPS; however, it experienced a drop in average recall compared to the baseline. Model A exhibited a substantial increase in parameter count to

5.78 M, with the FPS reduced to only 18.3, demonstrating that high-complexity modules are unsuitable for lightweight backbone integration.

Further comparison of the position-decoupled strategies highlights the necessity of maintaining a lightweight backbone. For all schemes introducing EMA into the backbone (Models E and F), the FPS remained below 27, indicating that computational redundancy in shallow network stages constitutes a major bottleneck for inference speed, without delivering proportional gains in overall *mAP* compared to Model D.

Model D (Ours) achieves the optimal trade-off among all compared methods. By adopting an asymmetric strategy that integrates SE into the backbone and CBAM into the detection head, the proposed model attains an excellent balance, achieving an average *mAP*@0.5 of 97.07% and an average Recall of 94.18%, while maintaining a stable real-time inference speed of 62.1 FPS. These results demonstrate that the proposed model effectively mitigates missed detections of subtle features in fixed industrial scenarios while satisfying real-time performance requirements.

However, a single-factor ANOVA was conducted for the six models to assess whether the differences in performance (*P*, *R*, *mAP*@0.5, and *mAP*@0.5:0.95) were statistically significant. The results showed that the differences in these metrics were not significant ($p > 0.05$), suggesting that the models' performance variations could largely be attributed to randomness in the training process rather than any structural changes.

Grad-CAM visualization analysis: To more intuitively reveal the underlying mechanism behind the performance improvements, Grad-CAM was employed to visualize and compare the feature extraction layers of the baseline model and Model D, as shown in Figure 5. The heatmaps of the baseline model exhibit relatively dispersed activation regions, with several high-response areas incorrectly located at the egg tray edges or specular reflection points, indicating susceptibility to background noise and insufficient focus on weak texture features. In contrast, Model D demonstrates a pronounced target-focused activation pattern: high-response regions are precisely concentrated on fertilization-related features (vascular textures) within the central region of the eggs, while background activations are effectively suppressed.

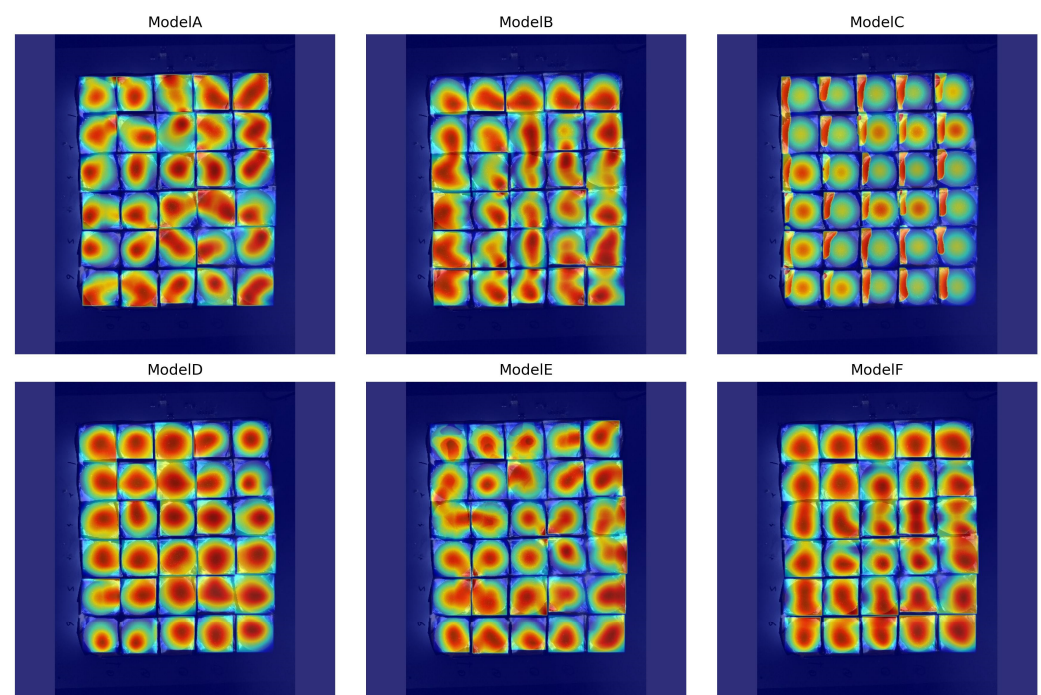


Figure 5. Heatmap visualization comparison of different models.

This visual evidence strongly validates the effectiveness of the proposed position-decoupled attention strategy. Specifically, the SE module embedded in the backbone effectively filters channel-wise interference caused by illumination noise, while the CBAM integrated into the detection head further enhances the spatial localization of critical texture features. The synergistic interaction between these two components enables accurate detection of weak-texture targets under complex illumination conditions.

In summary, Model D (Ours) achieves the optimal trade-off among all compared methods. By adopting an asymmetric strategy that integrates SE into the backbone and CBAM into the detection head, the proposed model attains an excellent balance, achieving an average $mAP@0.5$ of 97.07% and an average Recall of 94.18%, while maintaining a stable real-time inference speed of 62.1 FPS. These results demonstrate that the proposed model effectively mitigates missed detections of subtle features in fixed industrial scenarios while satisfying real-time performance requirements.

3.2. Robustness Analysis Under Different Illumination Intensities

Although the ablation experiments were conducted under controlled conditions, in practical industrial production, unavoidable disturbances such as light source aging (leading to illumination attenuation) and voltage fluctuations (resulting in instantaneous overexposure) are common. To evaluate the adaptability of Model D under non-ideal operating conditions, standard illumination (360 lumens) was used as the reference, and the model performance under extreme low-light (180 lumens) and high-light (540 lumens) conditions was comparatively analyzed. The detailed experimental results are presented in Table 3.

Table 3. Performance comparison under different illumination intensities.

Light Condition (Lumens)	Model	$P/\%$	$R/\%$	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$
180	Baseline	91.6	91.9	95.6	76.1
	Ours	87.6	97.4	96.6	77.8
360	Baseline	84.1	95.4	96.6	82.4
	Ours	87	91.5	97.4	82.6
540	Baseline	92.5	96.4	95.2	82
	Ours	89	97.7	97.8	77.6

The experimental results demonstrate that Model D (Ours) exhibits outstanding feature recovery capability and strong resistance to interference under extreme illumination conditions. Under low-light conditions at 180 lumens, benefiting from the amplification of critical channel weights by the SE module embedded in the backbone, the model effectively overcomes texture loss caused by dark-region noise, significantly improving the recall from 91.9% for the baseline model to 97.4%.

Conversely, under high-illumination conditions at 540 lumens, the CBAM deployed in the detection head suppresses background overexposure interference while emphasizing edge-related features, enabling the model to achieve an $mAP@0.5$ of 97.8%. However, severe illumination fluctuations lead to a slight decrease in $mAP@0.5-0.95$ compared to standard operating conditions. This decrease is primarily due to overexposure caused by excessive illumination, which results in the loss of important texture and edge details, affecting the model's ability to precisely localize objects. Despite this, the model maintains consistently high stability in the core binary recognition metrics (Recall and $mAP@0.5$), demonstrating that the proposed decoupled architecture remains effective even under challenging con-

ditions. These results strongly confirm that the proposed decoupled architecture exhibits robust practical performance in complex industrial environments.

3.3. PDA-YOLOv8 Training Performance Analysis

3.3.1. Precision–Recall Curve Analysis

The Precision–Recall (PR) curve provides a comprehensive representation of the trade-off between detection precision and recall under different confidence thresholds. Figure 6 illustrates the PR curves of Model D for each category, along with its overall detection performance on the training set.

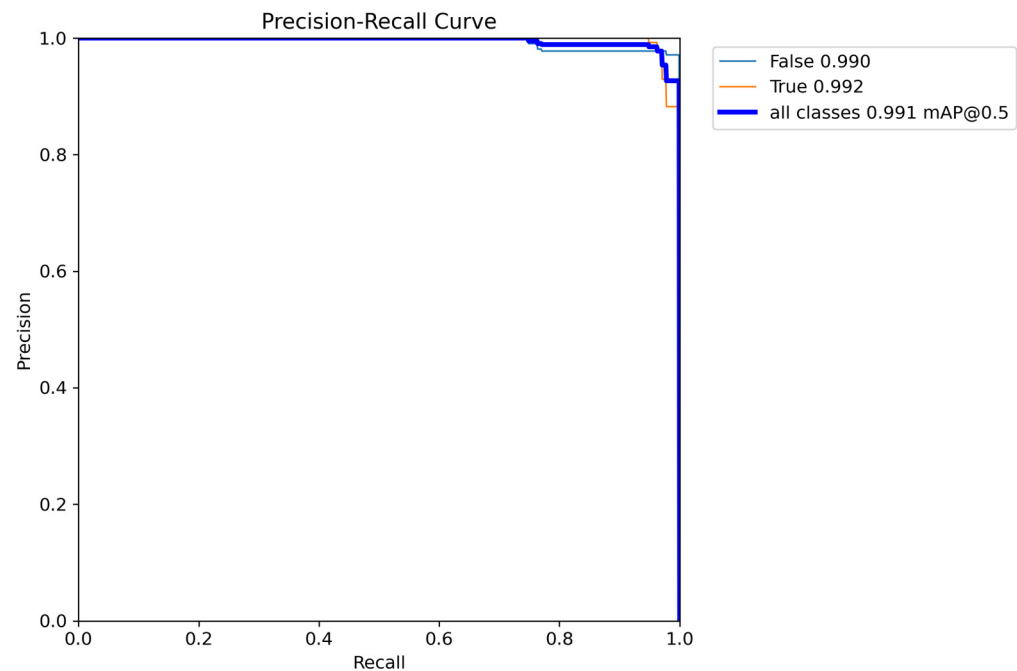


Figure 6. Precision–Recall (PR) curves of the proposed PDA-YOLO model for fertilized eggs, unfertilized eggs, and overall detection performance.

As shown in the figure, the PR curves for both fertilized and unfertilized egg categories are tightly clustered toward the upper-right corner of the coordinate space, exhibiting smooth and stable trends. Across a wide recall range (Recall < 0.95), the model maintains a precision above 0.99, indicating an extremely low false-positive rate under most confidence threshold settings.

Further category-wise comparison reveals that the average precision (AP) for fertilized eggs reaches 0.992, which is slightly higher than that for unfertilized eggs (0.990). This result suggests that the model demonstrates stronger and more stable recognition capability for fertilized features primarily characterized by subtle vascular textures. The PR curve shapes for the two categories are highly consistent, with no evident performance divergence, indicating the absence of significant classification bias.

In addition, a moderate decline in precision is observed in the extreme recall region approaching 1.0, which is mainly attributed to the inclusion of low-confidence predictions in the evaluation and represents a common phenomenon in object detection tasks. Overall, Model D achieves an $mAP@0.5$ of 0.991, further confirming its high stability and reliability in fertilized egg detection tasks.

3.3.2. Confusion Matrix Analysis

To further analyze the misclassification relationships among different categories during practical prediction, a confusion matrix was constructed based on the detection results

of Model D on the test set, as shown in Figure 7. The confusion matrix includes three classes—fertilized eggs, unfertilized eggs, and background—to comprehensively reflect the model’s discrimination capability between target and non-target regions.

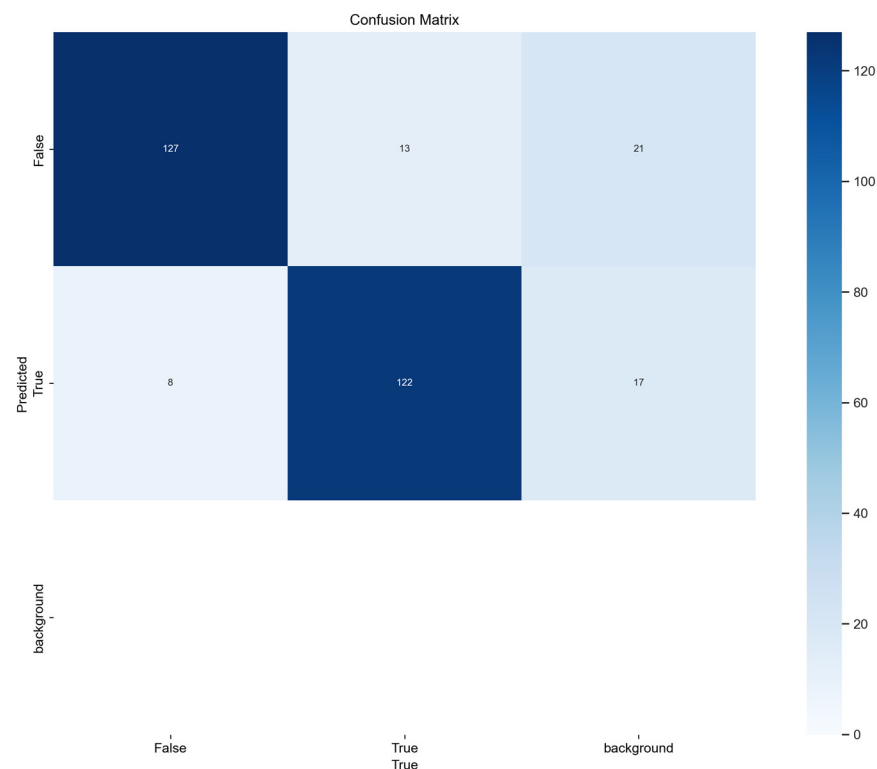


Figure 7. Confusion matrix of the proposed PDA-YOLO model across three categories.

As indicated by the diagonal elements, the model achieves high correct classification rates for both fertilized and unfertilized eggs. Specifically, 127 unfertilized egg samples and 122 fertilized egg samples are correctly identified, demonstrating a relatively balanced discrimination capability across the two key target categories.

Most misclassified samples originate from confusion between target categories and the background, rather than direct confusion between fertilized and unfertilized eggs. For instance, some samples with ground-truth labels of fertilized or unfertilized eggs are misclassified as background, which mainly occurs in regions near egg boundaries with transmissive effects or under strong local specular interference. In contrast, the proportions of fertilized eggs misclassified as unfertilized eggs, and vice versa, are relatively low, indicating strong stability in the model’s core category discrimination.

From an application perspective, this error distribution is considered reasonable. In industrial incubation scenarios, prioritizing accurate identification of fertilized eggs while assigning uncertain regions to the background class helps reduce the risk of incorrect decisions caused by noise interference. These results further demonstrate that the proposed position-decoupled attention mechanism effectively suppresses interference from background regions and enhances the reliability of target discrimination under complex imaging conditions.

3.4. Comparison Experiment

To comprehensively evaluate the performance and practical value of the proposed model (PDA-YOLO), comparative experiments were conducted against several mainstream state-of-the-art (SOTA) object detection models. The compared models range from classical lightweight networks (YOLOv5nu, YOLOv7-tiny) to the most recent iterations (YOLOv8n

through YOLOv12n), and the detailed quantitative results are summarized in Table 4. It is worth noting that to rigorously demonstrate the robustness of our model, the metrics reported for PDA-YOLO in Table 4 represent the average of 10 independent runs, whereas the baseline and other SOTA models report standard single-run best results.

Table 4. Performance comparison with state-of-the-art methods.

Model	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	Params/MB	Flops/G	FPS/(f.s ^{−1})
YOLOv5nu	91.4	94.0	94.8	81.0	2.51	7.2	80.6
YOLOv7tiny	82.6	90.0	98.3	82.3	6.02	13.2	31.1
YOLOv8n	84.1	95.4	96.6	82.4	3.01	8.2	78.5
YOLOv9t	87.9	88.1	95.3	81.5	2.01	7.8	26.2
YOLOv10n	85.5	93.1	93.8	79.4	2.71	8.4	270.3
YOLOv11n	90.7	89.0	96.2	82.2	2.59	6.4	19.8
YOLOv12n	95.7	91.9	97.1	83.5	2.57	6.5	66.7
ours	91.5	94.2	97.1	82.0	3.03	8.2	62.1

As the comparative analysis in Table 4 reveals, PDA-YOLO achieves the optimal trade-off among detection accuracy, inference speed, and model complexity. Compared to the baseline YOLOv8n, our model improves the core metric *mAP@0.5* from 96.6% to 97.07% with only a negligible increase in parameters (from 3.01 M to 3.03 M) and FLOPs. Although the integration of the decoupled attention mechanisms inevitably causes a reasonable decrease in pure inference speed (from 78.5 FPS to 62.1 FPS), it still successfully satisfies the stringent real-time requirement of industrial incubation lines (>60 FPS).

When compared to other architectures, our model demonstrates distinct advantages for this specific scenario. While YOLOv7-tiny achieves a marginally higher *mAP@0.5* of 98.3%, its parameter count (6.02 M) and FLOPs (13.2 G) are approximately double those of our model, resulting in a suboptimal inference speed of 31.1 FPS, which falls completely short of real-time processing needs. Conversely, the latest YOLOv10n exhibits an exceptional speed of 270.3 FPS due to its NMS-free architecture; however, it suffers a severe accuracy degradation (*mAP@0.5* drops to 93.8%), rendering it unsuitable for high-precision egg detection tasks. Furthermore, even when compared to the highly competitive YOLOv12n (*mAP@0.5* of 97.1%, 66.7 FPS), PDA-YOLO achieves comparable peak accuracy under a much more rigorous 10-run average evaluation protocol, underscoring its superior stability and robust feature extraction capability.

In summary, with a lightweight architecture of only 3.03 M parameters, the proposed model successfully maintains a real-time speed of 62.1 FPS while delivering highly stable and accurate detection performance, proving it to be the most outstanding holistic solution for this task.

4. Discussion

This study proposes an integrated solution for early fertilized egg detection, combining a high-throughput transmissive imaging system with a position-decoupled attention object detection algorithm. Although the experimental results demonstrate the superiority of PDA-YOLO, an in-depth comparative analysis with existing studies and an objective assessment of its limitations are necessary to accurately evaluate its industrial value.

4.1. In-Depth Comparative Analysis with Existing Studies

Previous computer vision research in poultry incubation has predominantly focused on single-egg detection or idealized lighting environments. For instance, although Li et al. and Geng et al. achieved extremely high classification accuracies (>98%) on poultry eggs, their methods lacked systematic evaluations of high-throughput batch processing (e.g., detecting 30 eggs simultaneously) and did not address the inference speed bottlenecks required for industrial conveyor belts.

Our comparative analysis with current SOTA models further elucidates this challenge. The empirical results indicate that the proposed position-decoupled attention strategy provides an effective balance between a lightweight architecture and feature extraction robustness. This demonstrates that for the specific characteristics. Compared to the standard YOLOv8n, PDA-YOLO can more stably extract faint vascular textures from complex backgrounds. Furthermore, in comparison with existing SOTA models, different models exhibited varying emphases: while the NMS-free YOLOv10n possesses an extremely high inference speed (270.3 FPS), its detection accuracy ($mAP@0.5$ of 93.8%) struggles to meet high-precision recognition requirements. Conversely, YOLOv7-tiny achieves higher accuracy but, constrained by a larger number of parameters, its inference speed drops to 31.1 FPS. In contrast, PDA-YOLO achieved an average $mAP@0.5$ of 97.07% over 10 strictly independent runs while maintaining an inference speed of 62.1 FPS, realizing a reasonable trade-off between precision and speed. This demonstrates that for the specific characteristics of industrial detection tasks, strategically planning the deployment positions of attention mechanisms helps enhance the comprehensive application potential of the model in real-world scenarios [30,35]. Additionally, Grad-CAM visualizations confirm that the decoupled attention strategy indicates improved attention to embryonic vascular textures, helping the model focus on relevant features while suppressing background interference from egg tray edges and reflections.

4.2. Limitations and Applicability Under Different Imaging Conditions

Despite the excellent performance of the model under varying illumination, the current dataset and experimental conditions still present certain limitations.

First, regarding its generalization capability to other types of hatching eggs, the current dataset is limited to Hy-Line White eggs. White-shelled eggs possess high light transmittance, making internal embryonic features relatively distinct under strong transmissive light. However, when applied to brown-shelled eggs (e.g., Hy-Line Brown) or thick-shelled waterfowl eggs (e.g., duck and goose eggs), the significantly reduced transmittance may lead to a substantial drop in image contrast and generate more severe dark-region noise. Additionally, speckled eggs (e.g., quail eggs) introduce shell texture noise that can easily be misclassified as early blood vessels. The current model has not been trained to decouple shell pigment noise from internal biological features, which restricts its direct generalization across different poultry breeds. Our robustness tests under varying light intensities (180.360.540 lumens) indicate that the model effectively mitigates optoelectrical interference that often leads to false negatives in industrial settings. At 60 h of incubation, corresponding to early embryonic development, the model achieves an $mAP@0.5$ of 97.4% and a Recall of 91.5%, demonstrating its ability to reliably detect embryonic features at early developmental stages. While previous automated systems often struggled with lighting fluctuations, our decoupled strategy maintains stable performance even under fluctuating illumination conditions, ensuring robust detection of fertilized eggs at the early embryonic stage.

Second, concerning its applicability under diverse imaging conditions, our image acquisition system relies on a customized darkroom to eliminate ambient light interference

and stabilize the samples. In real-world hatchery deployments, such image acquisition system designs will encounter several challenges: factors such as ambient light leakage from factory windows, continuous dust accumulation on the camera lens, and motion blur caused by conveyor belt vibrations were not systematically simulated in the current dataset. These dynamic noise sources could potentially weaken the model's feature extraction capability, leading to an increased false negative rate for viable embryos.

5. Conclusions

This study developed a deep learning-based, high-throughput automated early detection system for fertilized eggs—PDA-YOLO—successfully addressing the inefficiency and subjectivity of traditional candling methods.

The core contribution of this paper lies in proposing a novel position-decoupled attention strategy. By selectively embedding the lightweight SE module into the backbone network to amplify faint feature responses in low-contrast regions, and deploying the spatially aware CBAM in the detection head to suppress background clutter, we designed a model architecture tailored specifically for complex transmissive imaging scenarios. This architecture fundamentally resolves the conflict between maintaining lightweight computational efficiency and enhancing feature extraction robustness, achieving an optimal trade-off between high detection accuracy and industrial real-time processing speed under fluctuating illumination conditions.

To further promote the commercial application of this technology, future research will focus on two key directions. First, we plan to significantly expand the scale and diversity of the dataset to include samples of multiple breeds and shell colors (e.g., brown and speckled eggs), thereby improving the model's cross-species generalization capabilities. Second, the research focus will shift towards real-world deployment optimization, which includes integrating the model with edge computing devices and developing algorithms to compensate for motion blur and ambient light leakage on high-speed industrial conveyor belts, ultimately transforming this laboratory prototype into a plug-and-play intelligent solution for the global poultry industry.

Author Contributions: Conceptualization, F.L. and D.L.; methodology, Y.Z. and Z.S.; software, Y.Z.; validation, Y.Z.; formal analysis, Y.Z.; investigation, Y.Z.; data curation, Y.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, D.L., W.L. and F.L.; visualization, Y.Z.; supervision, Z.S.; project administration, D.L.; resources, G.Y. and H.P.; funding acquisition, F.L. and D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key R&D Program of Shandong Province, China (2022TZXD0017); the Shandong Provincial Natural Science Foundation (ZR2023QC120); the China Agriculture Research System (CARS-40); the Taishan Industry Experts Program (TSCX202312057); and Jinan City's "20 Measures" for Universities and Colleges (202534099).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the privacy policy of the organization.

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

References

1. Araújo, I.; Leandro, N.; Mesquita, M.; Café, M.; Mello, H.; Gonzales, E. Effect of incubator type and broiler breeder age on hatchability and chick quality. *Braz. J. Poult. Sci.* **2016**, *18*, 17–25. [[CrossRef](#)]
2. Boleli, I.; Morita, V.; Matos, J., Jr.; Thimotheo, M.; Almeida, V. Poultry egg incubation: Integrating and optimizing production efficiency. *Braz. J. Poult. Sci.* **2016**, *18*, 1–16. [[CrossRef](#)]
3. Nasri, H.; van den Brand, H.; Najjar, T.; Bouzouaia, M. Egg storage and breeder age impact on egg quality and embryo development. *J. Anim. Physiol. Anim. Nutr.* **2020**, *104*, 257–268. [[CrossRef](#)]

4. Avşar, K.O.; Uçar, A.; Özlü, S.; Elibol, O. Effect of high eggshell temperature during the early period of incubation on hatchability, hatch time, residual yolk, and first-week broiler performance. *J. Appl. Poult. Res.* **2022**, *31*, 100197. [\[CrossRef\]](#)
5. Franzo, G.; Swart, W.; Ugalde, M.A.; Cotta, H.; Lecoupeur, M.; Boyer, W.; Koutoulis, K.; Cecchinato, M. Impact of rotten eggs on hatchery performances: A multicentric study. *Animals* **2020**, *10*, 1725. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Hou, Z.; Cai, K.; Xiao, G.; Fang, Q.; Feng, Y.; Chen, H.; Yue, H.; Wang, Z.; Lin, Q. Early-stage fertilization detection in pigeon eggs during incubation using improved lightweight model. *Comput. Electron. Agric.* **2023**, *215*, 108422. [\[CrossRef\]](#)
7. Eghbalsaied, S.; Kues, W.A. Development of a transposon-based technology for transfection of day 0 chicken embryos. *Gene* **2020**, *730*, 144318. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Idahor, K.O. Avian reproduction. In *Animal Reproduction*; Bozkurt, Y., Bucak, M.N., Eds.; IntechOpen: London, UK, 2021. [\[CrossRef\]](#)
9. Ipek, A.; Sozcu, A. Comparison of hatching egg characteristics, embryo development, yolk absorption, hatch window, and hatchability of Pekin duck eggs of different weights. *Poult. Sci.* **2017**, *96*, 3593–3599. [\[CrossRef\]](#)
10. Li, Q.; Shao, Z.; Zhou, W.; Su, Q.; Wang, Q. MobileOne-YOLO: Improving the YOLOv7 network for the detection of unfertilized duck eggs and early duck embryo development—A novel approach. *Comput. Electron. Agric.* **2023**, *214*, 108316. [\[CrossRef\]](#)
11. Hosseini, S.A.; Banakar, A.; Minaei, S.; Mehdizadeh, S.A.; Torshizi, M.A.K.; Li, G. Determination of brown egg fertility by multi-directional visible/near-infrared spectroscopy and machine learning. *J. Appl. Poult. Res.* **2025**, *34*, 100534. [\[CrossRef\]](#)
12. Galli, R.; Preusse, G.; Schnabel, C.; Bartels, T.; Cramer, K.; Krautwald-Junghanns, M.-E.; Koch, E.; Steiner, G. Sexing of chicken eggs by fluorescence and Raman spectroscopy through the shell membrane. *PLoS ONE* **2018**, *13*, e0192554. [\[CrossRef\]](#)
13. Zhao, Q.; Ban, L.; Zheng, J.; Xu, G.; Ning, Z.; Qu, L. Potential use of spectroscopic techniques for assessing table eggs and hatching eggs. *World's Poult. Sci. J.* **2019**, *75*, 445–456. [\[CrossRef\]](#)
14. Ghaderi, M.; Mireei, S.A.; Masoumi, A.; Sedghi, M.; Nazeri, M. Chick embryo development assessment and fertility detection using pixel-wise hyperspectral image analysis and deep learning. *Poult. Sci.* **2025**, *104*, 106064. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Liu, L.; Ngadi, M.O. Detecting fertility and early embryo development of chicken eggs using near-infrared hyperspectral imaging. *Food Bioprocess. Technol.* **2013**, *6*, 2503–2513. [\[CrossRef\]](#)
16. Schreuder, J.; Niknafs, S.; Williams, P.; Roura, E.; Hoffman, L.C.; Cozzolino, D. Non-destructive prediction of fertility and sex in chicken eggs using the short wave near-infrared region. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2024**, *322*, 124716. [\[CrossRef\]](#)
17. Ghaderi, M.; Mireei, S.A.; Masoumi, A.; Sedghi, M.; Nazeri, M. Fertility detection of unincubated chicken eggs by hyperspectral transmission imaging in the Vis-SWNIR region. *Sci. Rep.* **2024**, *14*, 1289. [\[CrossRef\]](#)
18. Adegbenjo, A.O.; Liu, L.; Ngadi, M.O. An adaptive partial least-squares regression approach for classifying chicken egg fertility by hyperspectral imaging. *Sensors* **2024**, *24*, 1485. [\[CrossRef\]](#)
19. Wang, J.; Cao, R.; Wang, Q.; Ma, M. Nondestructive prediction of fertilization status and growth indicators of hatching eggs based on respiration. *Comput. Electron. Agric.* **2023**, *208*, 107779. [\[CrossRef\]](#)
20. Dong, J.; Lu, B.; He, K.; Li, B.; Zhao, B.; Tang, X. Assessment of hatching properties for identifying multiple duck eggs on the hatching tray using machine vision technique. *Comput. Electron. Agric.* **2021**, *184*, 106076. [\[CrossRef\]](#)
21. Huang, L.; He, A.; Zhai, M.; Wang, Y.; Bai, R.; Nie, X. A multi-feature fusion based on transfer learning for chicken embryo eggs classification. *Symmetry* **2019**, *11*, 606. [\[CrossRef\]](#)
22. Bouba, I.; Visser, B.; Kemp, B.; Rodenburg, T.B.; van den Brand, H. Predicting hatchability of layer breeders and identifying effects of animal-related and environmental factors. *Poult. Sci.* **2021**, *100*, 101394. [\[CrossRef\]](#)
23. Zhu, Z.H.; Ye, Z.F.; Tang, Y. Nondestructive identification for gender of chicken eggs based on GA-BPNN with double hidden layers. *J. Appl. Poult. Res.* **2021**, *30*, 100203. [\[CrossRef\]](#)
24. Geng, L.; Yan, T.; Xiao, Z.; Xi, J.; Li, Y. Hatching eggs classification based on deep learning. *Multimed. Tools Appl.* **2018**, *77*, 22071–22082. [\[CrossRef\]](#)
25. Çevik, K.K.; Koçer, H.E.; Boğa, M. Deep learning based egg fertility detection. *Vet. Sci.* **2022**, *9*, 574. [\[CrossRef\]](#)
26. Ultralytics. Ultralytics YOLOv8. Available online: <https://github.com/ultralytics/ultralytics> (accessed on 1 January 2025).
27. Bui, T.D.; Do Le, T.M. Ghost-Attention-YOLOv8: Enhancing rice leaf disease detection with lightweight feature extraction and advanced attention mechanisms. *AgriEngineering* **2025**, *7*, 93. [\[CrossRef\]](#)
28. Yang, C.; Huang, Q.; Yang, Q.; Li, L. An enhanced YOLOv8 algorithm for fish school recognition integrating SEAttention and LDConv. *Discov. Appl. Sci.* **2025**, *7*, 1201. [\[CrossRef\]](#)
29. Yang, G.; Wang, J.; Nie, Z.; Yang, H.; Yu, S. A lightweight YOLOv8 tomato detection algorithm combining feature enhancement and attention. *Agronomy* **2023**, *13*, 1824. [\[CrossRef\]](#)
30. Wei, Z.; Chang, M.; Zhong, Y. Fruit freshness detection based on YOLOv8 and SE attention mechanism. *Acad. J. Sci. Technol.* **2023**, *6*, 195–197. [\[CrossRef\]](#)
31. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141. [\[CrossRef\]](#)

32. Ronghua, Z.; Xue, B.; Jiangchuan, F. Crop pest target detection algorithm in complex scenes: YOLOv8-Extend. *Smart Agric.* **2024**, *6*, 49. [[CrossRef](#)]
33. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19. [[CrossRef](#)]
34. Li, X.; Cai, C.; Song, B.; Yang, Y. YOLOV8-MR: An improved lightweight YOLOv8 algorithm for tomato fruit detection. *IEEE Access* **2025**, *13*, 48120–48131. [[CrossRef](#)]
35. Yang, Q.; Gu, J.; Xiong, T.; Wang, Q.; Huang, J.; Xi, Y.; Shen, Z. RFA-YOLOv8: A robust tea bud detection model with adaptive illumination enhancement for complex orchard environments. *Agriculture* **2025**, *15*, 1982. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.