LIGHTWEIGHT DETECTION OF EGG SURFACE DEFECTS UNDER CONVEYOR-BELT CONDITIONS USING YOLOv11N-DAL

/

传送带条件下基于 YOLOv11n-DAL 的鸡蛋外观缺陷轻量化检测

Xianyong MENG[1,2)], Haoran GUO[1,2)], Jun YAN[1,2)], Qimiao WANG[1,2,*)], Qinyou SUN[1,2,*)]

[1)] School of Information Science and Engineering, Shandong Agricultural University, Taian 271018, China;

[2)] Key Laboratory of Huanghuaihai Smart Agricultural Technology, Ministry of Agriculture and Rural Affairs, Taian 271018, China;

** Corresponding author: E-mail: 15753837976@163.com;2024121207@sdau.edu.cn*

DOI: https://doi.org/10.35633/inmateh-79-77

***Keywords:** egg surface defects; lightweight detection; YOLOv11n-DAL; conveyor-belt sorting; machine vision*

ABSTRACT

Egg appearance defect detection in conveyor-belt environments is challenged by irregular defect morphology, illumination variations, and the limited computational resources of intelligent sorting equipment, making it difficult to simultaneously achieve high detection accuracy and lightweight deployment. In this study, a lightweight detection model, termed YOLOv11n-DAL, is proposed to improve egg defect detection performance while reducing computational complexity. The proposed model integrates C3k2_DCNv4, ADown, and Detect_LSDECD modules into the YOLOv11n framework to enhance defect feature representation, improve effective information retention during downsampling, and reduce computational redundancy in the prediction stage. A conveyor-belt egg image dataset containing three categories, namely damaged, dirty, and good, was constructed. The dataset consisted of 1,730 original images, and the training set was expanded to 4,152 images through data augmentation. Comparative experiments and ablation studies were conducted to validate the effectiveness of the proposed model. The experimental results demonstrated that YOLOv11n-DAL achieved 95.2% Precision, 94.5% Recall, 94.8% F1-score, 96.7% mAP@0.5, and 95.8% mAP@0.5:0.95. Compared with the original YOLOv11n model, YOLOv11n-DAL reduced the number of parameters from 2.6 M to 1.7 M, GFLOPs from 6.3 to 4.8, and model size from 5.2 MB to 3.8 MB, while maintaining superior detection performance. The results indicate that YOLOv11n-DAL achieves an effective balance between detection accuracy and computational efficiency, providing a lightweight and practical solution for intelligent egg sorting systems in conveyor-belt environments.

摘要

传送带环境下鸡蛋外观缺陷检测受到缺陷形态不规则、光照变化以及智能分拣设备计算资源有限等因素影响，难以同时满足检测精度和轻量化部署需求。本文提出一种轻量化检测模型 YOLOv11n-DAL，在降低计算复杂度的同时提升鸡蛋缺陷检测性能。该模型将 C3k2_DCNv4、ADown 和 Detect_LSDECD 融合至 YOLOv11n 框架中，以增强缺陷特征表达能力、提高下采样过程中的有效信息保留能力，并降低预测阶段的计算冗余。本文构建了包含 damaged、dirty 和 good 三类的传送带鸡蛋图像数据集，包括 1730 张原始图像，训练集经数据增强后扩展至 4152 张。通过对比实验和消融实验对模型性能进行验证。实验结果表明，YOLOv11n-DAL 模型的 Precision、Recall、F1-score、mAP@0.5 和 mAP@0.5:0.95 分别达到 95.2%、94.5%、94.8%、96.7%和 95.8%。相比 YOLOv11n，所提出模型将参数量由 2.6 M 降低至 1.7 M，GFLOPs 由 6.3 降低至 4.8，模型大小由 5.2 MB 降低至 3.8 MB，同时保持更高检测性能。研究结果表明，YOLOv11n-DAL 能够在检测精度和计算效率之间取得良好平衡，为传送带鸡蛋智能分拣系统提供了一种轻量化解决方案。

INTRODUCTION

Eggs are important poultry products with high nutritional and economic value, and their external quality directly affects commercial grading, storage stability, transportation safety, and consumer acceptance. During collection, conveying, grading, and packaging, eggs are easily affected by mechanical impact, collision, and environmental contamination, resulting in defects such as cracks, shell breakage, and surface dirt. These defects not only reduce product value but may also increase potential food safety risks. Therefore, rapid, accurate, and non-destructive external-defect detection is essential for automated egg sorting and intelligent quality control in modern poultry processing systems (*Gautron et al., 2021; So and Kim, 2022; Yang et al., 2023; Atwa et al., 2024; Yu et al., 2025*).

Traditional manual inspection methods are still widely used in some egg-processing facilities. However, manual inspection is highly dependent on operator experience and is affected by fatigue, subjective judgment, and low efficiency, making it difficult to satisfy the requirements of high-speed industrial sorting. With the development of intelligent agricultural equipment, machine vision and deep learning technologies have gradually become effective solutions for agricultural product quality evaluation due to their advantages of rapid image acquisition, objective decision-making, and automated operation. In recent years, deep learning-based visual methods have been successfully applied to egg defect classification, eggshell crack detection, freshness evaluation, and online quality monitoring (*Turkoglu, 2021; Botta et al., 2022; Chen et al., 2023; Huang et al., 2023; Zhao et al., 2023; Yin et al., 2024; Çengel et al., 2025; Wu et al., 2025*).

Among deep learning-based object detection algorithms, the YOLO series has attracted extensive attention in agricultural visual applications because of its excellent balance between detection accuracy and computational efficiency (*Kanna et al., 2024; Gupta et al., 2025*). Compared with image classification methods, object detection algorithms can simultaneously achieve target localization and category recognition, making them more suitable for continuous conveyor-belt inspection scenarios. Recent studies have demonstrated that improved YOLO models can effectively enhance detection performance under complex agricultural environments. *Tang et al., (2026),* proposed an improved RT-DETR model for defective egg detection in edge-computing environments, demonstrating the effectiveness of advanced detection architectures for intelligent egg sorting applications. *Ma et al., (2025),* developed YOLO-TRS based on YOLO11 for tomato fruit ripeness and stem detection, demonstrating the effectiveness of YOLO11-based lightweight models in agricultural vision tasks. *Shi et al., (2025),* proposed an improved YOLOv11n-based model for missing-seed detection and counting, further confirming the potential of lightweight networks for intelligent agricultural equipment deployment. In addition, *Bing et al., (2025),* introduced improved feature extraction strategies into YOLOv8s for apple leaf lesion detection under complex backgrounds, indicating that feature enhancement is important for small-object detection tasks.

However, practical application of deep learning-based egg inspection systems still faces several challenges. First, egg defects such as small cracks, irregular damaged regions, and weak dirty spots usually exhibit unclear boundaries and low contrast, making them easily confused with eggshell texture, illumination variation, and conveyor-belt background interference. Second, some existing methods rely on image classification or specialized imaging systems, which limits their application in low-cost and high-speed sorting lines. Although hyperspectral and other advanced imaging technologies can improve defect visibility, their high equipment cost and complex configuration restrict large-scale industrial deployment. Third, large-scale detection networks generally require considerable computational resources, making them difficult to deploy on embedded devices and edge-computing platforms. YOLO11 provides a lightweight backbone-neck-head structure with improved computational efficiency; however, its original architecture still has limitations for fine-grained egg external-defect detection, since redundant feature fusion and repeated prediction-head operations increase computational overhead without proportionally improving sensitivity to irregular defect boundaries *(Jocher et al., 2024*).

In recent years, feature-enhancement and computational-optimization strategies for lightweight object detection networks have provided new research directions for addressing these problems. Deformable convolution was first introduced by *Dai et al. (2017)* to achieve flexible kernel deformation by learning content-adaptive sampling offsets, and was later refined into the more efficient DCNv4 by *Xiong et al. (2024)*; DCNv4-based modules have already been applied to irregular-target detection tasks outside agriculture, such as steel surface defect detection (*Zhu et al., 2025*). For efficient downsampling, *Wang et al. (2024)* proposed the ADown module in YOLOv9, which has since been used in YOLOv11-based agricultural detection models to preserve small-target features under complex backgrounds (*Wang et al., 2026*). For the detection head, Detail-Enhanced Convolution (DEConv), proposed by *Chen et al. (2024)* for image dehazing, enhances fine-detail representation without additional inference cost when combined with a lightweight shared convolutional head. Although each of these modules has been individually validated, their joint application to egg external-defect detection under conveyor-belt conditions has not yet been reported.

Based on the above, this study proposes a lightweight detection model, YOLOv11n-DAL, for egg external-defect detection under conveyor-belt conditions, in which three modules are jointly designed to improve detection accuracy while reducing model parameters and computational cost. Specifically, DCNv4 is introduced into the C3k2 structure of YOLOv11n to construct the C3k2_DCNv4 module, enhancing the representation of irregular defect features; ADown replaces the original downsampling convolution, reducing model size while improving feature preservation; and Detect_LSDECD combines a lightweight shared convolutional structure with DEConv to reduce redundant prediction-head computation while enhancing detail sensitivity. Together, these modules achieve both improved accuracy and reduced model complexity, offering a feasible solution for resource-constrained conveyor-belt sorting equipment. A conveyor-belt egg image dataset containing the damaged, dirty, and good categories was constructed, and the proposed method was comprehensively evaluated.

MATERIALS AND METHODS

Data Collection and Dataset Construction

Egg images were collected from a conveyor belt located in Jining City, Shandong Province, China, from February 8 to February 10, 2026, using an iPhone 17 Pro Max. To reduce motion blur and improve image quality, the conveyor belt was temporarily stopped during image acquisition. A top-view imaging configuration was adopted, where the smartphone camera was positioned vertically above the conveyor belt and kept parallel to its surface. The shooting distance was approximately 10–60 cm, and the original image resolution was 3072 × 4096 pixels. The data collection environment involved a combination of natural light and indoor illumination conditions to simulate the lighting variations that may occur in practical egg inspection scenarios. Each egg was captured from only a single viewpoint to prevent potential data leakage caused by the same instance appearing across different subsets.

A total of 1730 original JPG images were retained. The category labels were determined according to the overall appearance characteristics of each egg. Eggs with cracks, shell breakage, or holes were labeled as damaged. Eggs exhibiting surface contamination, such as feces, blood stains, or mud stains, but without structural damage were labeled as dirty. Eggs without apparent defects were labeled as good. When an egg simultaneously exhibited both structural damage and surface contamination, it was consistently assigned to the damaged category, since structural defects represent more severe quality issues. Representative samples of the three categories are shown in Fig. 1.

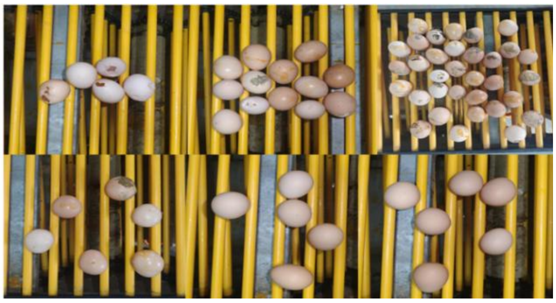


Fig. 1 - Representative images of the good, damaged, and dirty egg classes collected under conveyor-belt conditions

Dataset Annotation and Data Augmentation

The dataset was randomly divided into training, validation, and testing subsets at a ratio of 8:1:1, resulting in 1384, 173, and 173 images, respectively. Image annotation was manually performed using the Roboflow platform. Rectangular bounding boxes were generated to enclose the entire egg region. Data augmentation was applied only to the training set, including horizontal flipping, vertical flipping, ±90° rotation, ±10° random rotation, ±10° shearing, ±25% saturation adjustment, ±15% brightness adjustment, and noise injection with a maximum intensity of 2.24% pixels. For each original training image, three augmented images were generated, expanding the training set to 4152 images (Fig. 2).

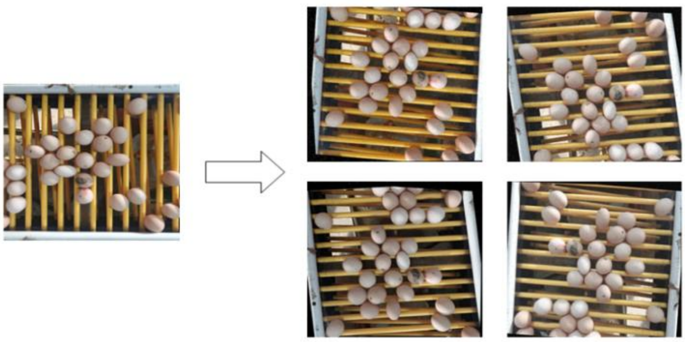


Fig. 2 - Examples of data augmentation effects

The final dataset contained 28,893 egg object instances in total. The numbers of instances belonging to the damaged, dirty, and good categories are summarized in Table 1.

Table 1

Distribution of annotated egg instances in the complete dataset.

Egg category	Annotated instances	Proportion (%)
damaged	8582	29.7
dirty	9449	32.7
good	10862	37.6
Total	28893	100.0

YOLOv11n-DAL Network Architecture and Improvement Methods

YOLOv11n was selected as the baseline because of its compact one-stage backbone–neck–head architecture and P3–P5 multi-scale prediction. To improve its sensitivity to small and irregular eggshell defects without increasing model complexity, YOLOv11n-DAL replaces the original C3k2 blocks with C3k2_DCNv4, adopts ADown for downsampling, and uses Detect_LSDECD as the prediction head. The overall architecture is shown in Fig. 3.

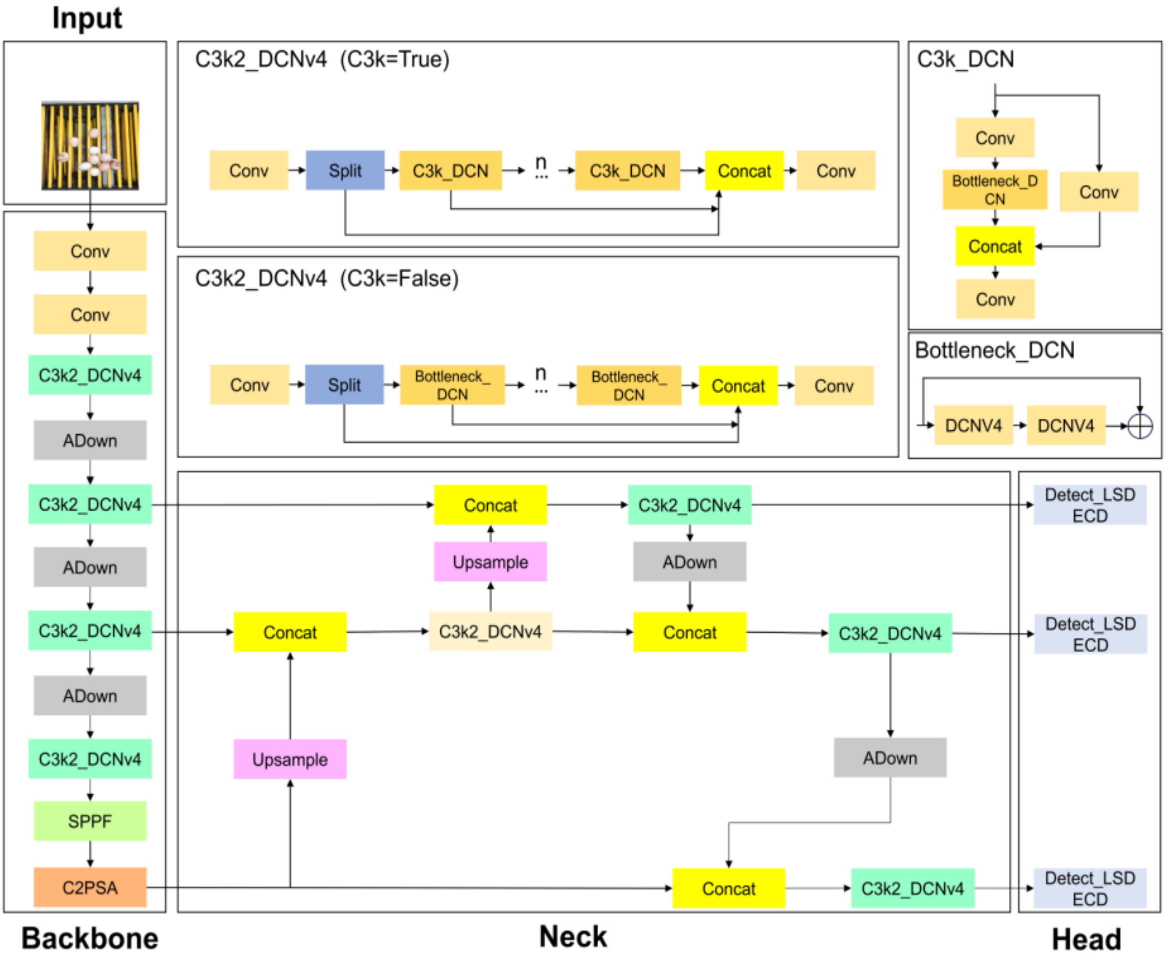


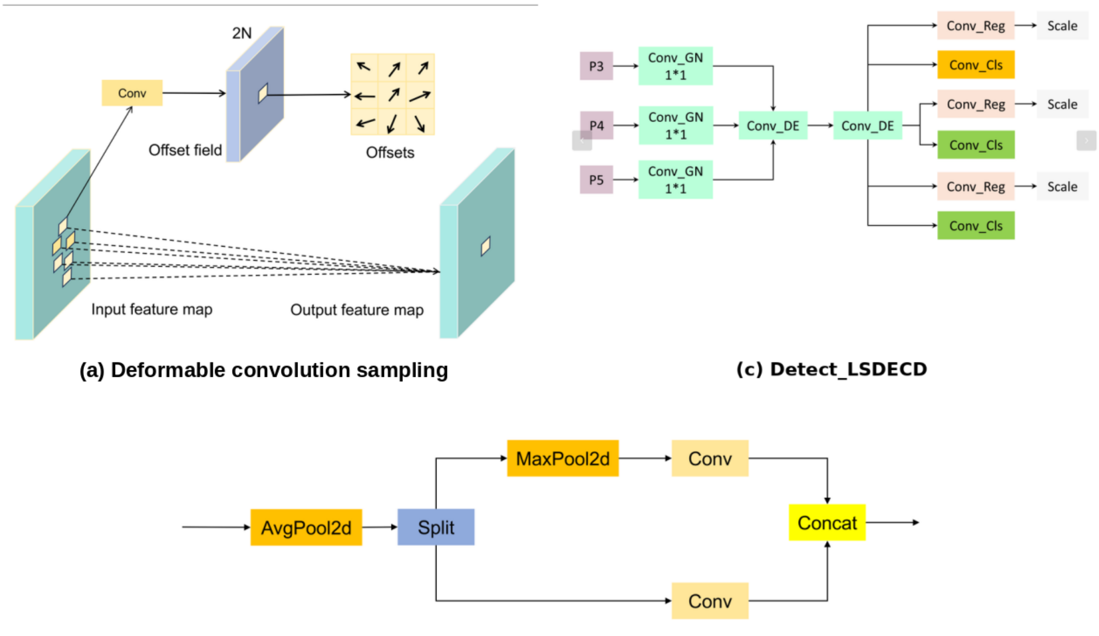
Fig. 3 - Overall architecture of the proposed YOLOv11n-DAL model for egg surface defect detection

C3k2_DCNv4 and ADown are deployed in the backbone and neck for adaptive feature extraction and efficient multi-scale transmission. Detect_LSDECD receives the P3, P4, and P5 feature maps and predicts the damaged, dirty, and good categories. The structures of the three improved modules are summarized in Fig. 4.

C3k2_DCNv4 Feature Extraction Module

C3k2_DCNv4 integrates deformable convolution into the C3k2 framework. As shown in Fig. 3, the input undergoes a 1×1 convolution, then splits into two branches.

One branch passes through n stacked C3k_DCN (or Bottleneck_DCN) blocks, while the other serves as a shortcut. The two branches are concatenated and fused by another 1×1 convolution. Each C3k_DCN block contains two DCNv4 layers with residual connections. DCN [Fig. 4(a)] learns offset fields to sample adaptively from irregular positions, improving the model's ability to handle geometric variations while maintaining efficiency.



(a) Deformable convolution sampling **(c) Detect_LSDECD**

(b) ADown

Fig. 4 - Structures of the improved modules in YOLOv11n-DAL
(a) Deformable convolution sampling; (b) ADown; (c) Detect_LSDECD

Unlike fixed-grid convolution, DCNv4 learns offsets according to target geometry:

$$y(p_0) = \sum_{k=1}^{K} w_k \; x(p_0 + p_k + \Delta p_k) \tag{1}$$

where x and y denote the input and output feature maps, p_0 is the current position, p_k is the regular sampling position, Δp_k is the learnable offset, and w_k is the convolution weight. This adaptive sampling is suitable for cracks, holes, broken-shell regions, and contaminants with irregular boundaries.

ADown Downsampling Module

ADown first applies average pooling and channel splitting. One branch uses a 3 × 3 stride-2 convolution, while the other uses max pooling followed by a 1 × 1 convolution; the two outputs are concatenated (Fig. 4(b)).

$$X_1, X_2 = \mathrm{Split}\big(\mathrm{AvgPool}(X)\big) \tag{2}$$

$$Y = \mathrm{Concat}(Y_1, Y_2) \tag{3}$$

where X is the input feature map, X_1 and X_2 are the two channel branches, $Y_1 = Conv_{3\times3,s=2}(X_1)$, $Y_2 = Conv_{1\times1}(MaxPool(X_2))$, and Y is the concatenated output. The dual-branch strategy combines average and maximum responses, helping preserve edge and texture information during downsampling.

Detect_LSDECD Detection Head

Detect_LSDECD first maps multi-scale features to a unified hidden dimension and then applies shared detail-enhanced convolutions, reducing repeated operations across detection layers (Fig. 4(c)).

$$F_i' = H_{\mathrm{shared}}\big(\mathrm{Conv}_{1\times1}(F_i)\big) \tag{4}$$

where F_i is the input of the i-th scale, $Conv_{1\times1}$ performs channel adjustment, H_{shared} denotes the shared feature extractor, and F_i' is the refined feature map.

$$B_i = \mathrm{DFL}(R_i), \quad C_i = \sigma(S_i) \tag{5}$$

where B_i and C_i are the predicted box and class probability, R_i and S_i denote the regression and classification outputs obtained from F_i', DFL denotes distribution-based box regression, and σ is the sigmoid function. Separate branches retain prediction quality while the shared extractor reduces head redundancy.

Training Environment and Parameter Settings

All models were trained and evaluated under the same experimental conditions to ensure a fair comparison. The hardware platform was equipped with an NVIDIA A10G GPU, and the software environment included Python 3.10.16, PyTorch 2.2.2, CUDA 12.1, and Ultralytics 8.3.9. The same training strategy and hyperparameter settings were applied to YOLOv11n-DAL and all comparison models. The main training hyperparameters are summarized in Table 2.

Table 2

Main training hyperparameter settings

Parameter	Value
Image size	640 × 640
Epochs	200
Batch size	16
Optimizer	SGD
Initial learning rate	0.01
Workers	10
close_mosaic	10

Model Performance Evaluation Indicators

Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95 were used to evaluate detection accuracy. Parameters, GFLOPs, and model size reflected model complexity. Inference speed was measured in FPS on a single GPU with batch size 1 and input size 640×640, including preprocessing, forward pass, and post-processing, Precision, Recall, and F1-score were calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \tag{6}$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \tag{7}$$

where TP, FP, and FN denote true positives, false positives, and false negatives, respectively; P and R in Eq. (8) denote Precision and Recall.

$$F_1 = \frac{2PR}{P + R} \times 100\% \tag{8}$$

The mean Average Precision over N egg categories was calculated as:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \times 100\% \tag{9}$$

where AP_i is the Average Precision of the i-th category. $mAP@0.5$ uses an IoU threshold of 0.5, whereas $mAP@0.5{:}0.95$ averages thresholds from 0.5 to 0.95 in steps of 0.05.

Human Baseline Validation on the Test Set

To further evaluate the performance of YOLOv11n-DAL in practical egg appearance defect detection scenarios, a comparative experiment between manual inspection and model-based detection was conducted using 173 original images from the test set. Two inspectors with more than three years of experience in egg quality inspection were invited to participate in the experiment.

The two inspectors independently performed defect classification on the egg targets in each test image without communication or discussion. Each egg sample was categorized into one of three classes: damaged, dirty, or good. Precision, recall, and F1-score were calculated separately based on the independent annotations of the two inspectors, and the arithmetic mean of these metrics was used as the baseline performance of manual inspection.

Subsequently, YOLOv11n-DAL was evaluated on the same test set using identical evaluation metrics and calculation procedures. By comparing the Precision, Recall, and F1-score between the proposed model and the manual inspection baseline, the performance differences between YOLOv11n-DAL and experienced quality inspectors in egg appearance defect recognition were analyzed.

RESULTS

YOLOv11n-DAL Model Performance Analysis

To comprehensively evaluate the detection performance of the improved model, Precision, Recall, mAP@50, and mAP@50:95 were adopted as evaluation metrics to analyze the model's performance across different categories. Table 3 presents the per-class results. As shown, the mAP@50 values for the damaged, dirty, and good categories reached 97.9%, 95.2%, and 97.0%, respectively. The overall Precision, Recall, mAP@50, and mAP@50:95 of the model were 95.2%, 94.5%, 96.7%, and 95.8%, respectively.

Table 3

Performance comparison of the YOLOv11n-DAL model on different egg categories

Class	P (%)	R (%)	F1 (%)	mAP50 (%)	mAP50–95 (%)
damaged	96.6	95.3	96.0	97.9	97.7
dirty	94.6	91.4	93.0	95.2	94.3
good	95.2	97.0	95.6	97.0	95.3
all	95.2	94.5	94.8	96.7	95.8

The normalized confusion matrix and visualization results are presented in Fig. 5. Fig. 5(a) compares heatmaps generated by YOLOv11n and YOLOv11n-DAL. The proposed model produces more concentrated activations on egg and defect regions, with weaker responses on conveyor rollers. Fig. 5(c) also demonstrates more complete target coverage for densely distributed, partially occluded, and visually ambiguous eggs, corroborating the quantitative improvements.

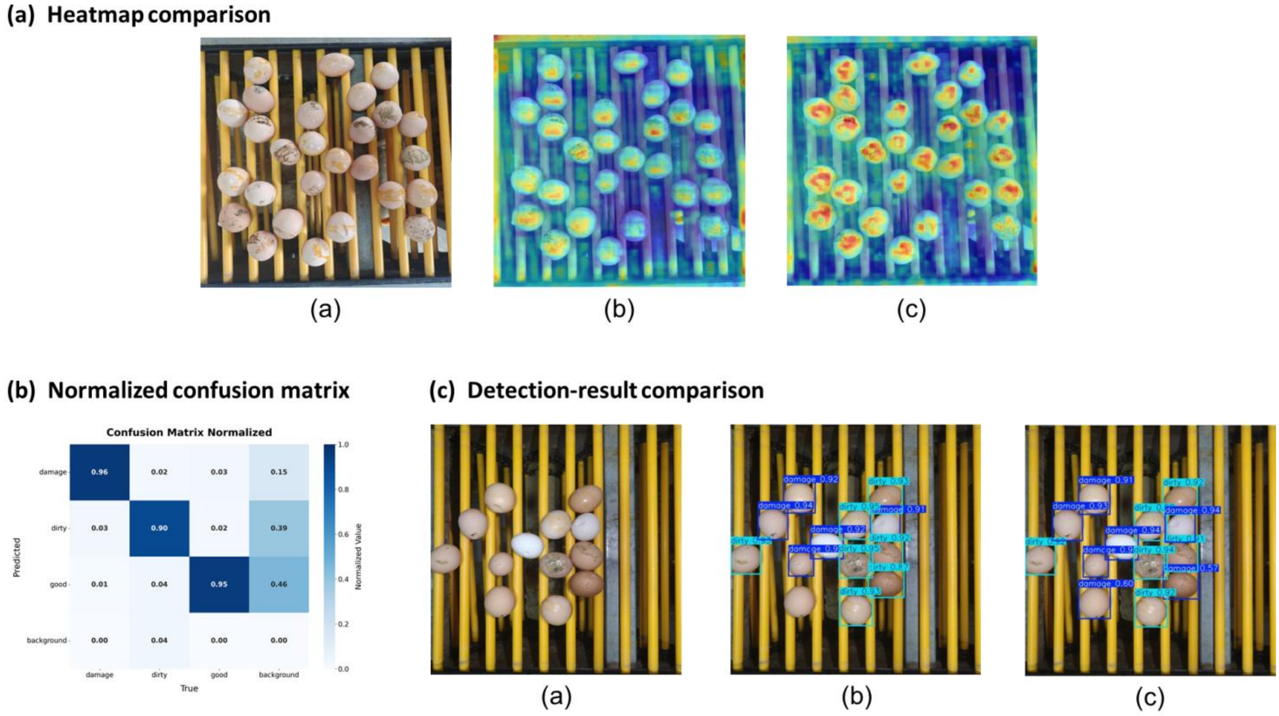


Fig. 5 - Visualization comparison between YOLOv11n and YOLOv11n-DAL: heatmaps (upper), normalized confusion matrix of YOLOv11n-DAL (lower left), and detection-result comparison (lower right)

The row-normalized confusion matrix in Fig. 5(b) shows class-wise recall of 96%, 90%, and 95% for damaged, dirty, and good, respectively. The dirty category remains comparatively more challenging, as slight stains, natural eggshell spots, shadows, and illumination-induced dark regions can appear visually similar. Some background false positives persist, indicating that weak defects and complex textures are the primary sources of missed or incorrect detections.

Ablation Experiments

To verify the effectiveness of the improved modules, ablation experiments were conducted using YOLOv11n as the baseline model. The effects of ADown, C3k2_DCNv4, and Detect_LSDECD were evaluated separately and jointly, and the results are shown in Table 4.

Table 4

Ablation study results of YOLOv11n-DAL

Exp.	ADown	DCNv4	LSDECD	P (%)	R (%)	F1 (%)	mAP50 (%)	mAP50–95 (%)	Params (M)	GFLOPs	Size (MB)
1	×	×	×	93.5	93.5	93.5	95.8	94.8	2.6	6.3	5.2
2	√	×	×	95.8	92.6	94.2	96.1	94.8	2.0	5.3	4.1
3	×	√	×	95.5	92.2	93.8	96.3	95.4	2.5	6.2	5.1
4	×	×	√	95.5	92.9	94.2	96.9	95.6	2.3	6.0	5.0
5	√	√	×	94.9	93.1	94.0	96.2	95.3	2.0	5.1	4.1
6	√	×	√	93.4	95.3	94.3	96.9	96.0	1.8	5.0	4.1
7	×	√	√	95.5	93.8	94.6	96.4	95.4	2.1	5.8	4.7
8	√	√	√	95.2	94.5	94.8	96.7	95.8	1.7	4.8	3.8

Note: AD denotes ADown, DCNv4 denotes C3k2_DCNv4, and LSDECD denotes Detect_LSDECD; √ indicates that a module is included, whereas × indicates that it is excluded; bold values indicate the best result for each metric.

Compared with the baseline model, incorporating ADown reduced the number of parameters, computational complexity, and model size from 2.6 M, 6.3 GFLOPs, and 5.2 MB to 2.0 M, 5.3 GFLOPs, and 4.1 MB, respectively, while maintaining comparable detection performance. When introduced individually, C3k2_DCNv4 and Detect_LSDECD increased mAP@0.5:0.95 from 94.8% to 95.4% and 95.6%, respectively, indicating their effectiveness in enhancing feature extraction and localization performance.

Furthermore, the combinations of ADown + C3k2_DCNv4 (Exp. 5), ADown + Detect_LSDECD (Exp. 6), and C3k2_DCNv4 + Detect_LSDECD (Exp. 7) were evaluated to investigate the complementary effects of the proposed modules. By integrating all three modules, the resulting YOLOv11n-DAL achieved the best overall performance, obtaining a precision of 95.2%, recall of 94.5%, F1-score of 94.8%, mAP@0.5 of 96.7%, and mAP@0.5:0.95 of 95.8%. Meanwhile, the model size was reduced to 1.7 M parameters, 4.8 GFLOPs, and 3.8 MB. These results indicate that the proposed modules simultaneously enhance detection accuracy while reducing computational cost and model size, demonstrating the effectiveness of YOLOv11n-DAL for lightweight object detection.

Comparison with Different Detection Models

To further evaluate the comprehensive performance of YOLOv11n-DAL, several representative detection models were selected for comparison, including YOLOv6, YOLOv8, YOLOv10, YOLOv11n, YOLO12n, YOLO26n, Hyper-YOLO, and RT-DETR-R18. All models were trained and tested under the same experimental conditions, and the results are shown in Table 5.

Table 5

Performance comparison results of different detection models

Model	P (%)	R (%)	F1 (%)	mAP50 (%)	mAP50–95 (%)	Params (M)	GFLOPs	Size (MB)	FPS
YOLOv6n	94.2	91.3	92.7	95.9	95.2	4.2	11.4	8.2	146.5
YOLOv8n	93.0	93.9	93.4	96.2	95.2	2.7	6.8	5.4	138.5
YOLOv10n	94.0	91.4	92.7	95.3	94.1	2.3	6.5	5.5	127.9
YOLOv11n	93.5	93.5	93.5	95.8	94.8	2.6	6.3	5.2	96.7
YOLO12n	95.0	94.0	94.5	96.7	95.8	2.5	6.3	5.3	78.5
YOLO26n	92.4	91.3	91.8	95.5	93.6	2.5	5.2	5.2	89.6
Hyper-YOLO	96.3	92.4	94.3	96.6	95.7	3.6	9.5	7.3	88.9
RT-DETR-R18	95.2	95.4	95.3	95.1	93.9	19.9	56.9	307.3	100.1
YOLOv11n-DAL	95.2	94.5	94.8	96.7	95.8	1.7	4.8	3.8	61.0

Note: Model size refers to the saved weight-file size generated by the corresponding implementation.

YOLOv11n-DAL achieves a precision of 95.2%, recall of 94.5%, F1-score of 94.8%, mAP@0.5 of 96.7%, and mAP@0.5:0.95 of 95.8%. It achieves the highest mAP@0.5 (96.7%) among all methods, tied with YOLO12n, while surpassing the baseline YOLOv11n by 0.7% in mAP@0.5 and 0.9% in mAP@0.5:0.95.

More importantly, these gains are achieved with only 1.7 M parameters, 4.8 GFLOPs, and 3.8 MB model size—a 34.6% reduction in parameters and 23.8% reduction in GFLOPs compared to the baseline. In contrast, RT-DETR-R18 requires 19.9 M parameters and 56.9 GFLOPs, and Hyper-YOLO requires 3.6 M parameters and 9.5 GFLOPs.

Inference speed is 61.0 FPS, lower than the baseline 96.7 FPS due to DCNv4-based feature fusion, but still well above the 30 FPS real-time requirement for industrial inspection. The compact model size (3.8 MB) also alleviates edge-device memory constraints, making YOLOv11n-DAL a more practical solution for edge deployment.

Comparison with Human Inspectors

To further evaluate the practical applicability of the proposed YOLOv11n-DAL, its performance was compared with that of human inspectors, as shown in Table 6.

Table 6

Detection performance of YOLOv11n-DAL vs. human inspectors on the test set

Method	P (%)	R (%)	F1 (%)
Inspector 1	91.3	91.3	91.0
Inspector 2	88.6	88.4	88.0
Average of manual inspectors	89.9	89.8	89.5
YOLOv11n-DAL	95.2	94.5	94.8

As shown in Table 6, the two inspectors achieved precision values of 91.3% and 88.6%, recall values of 91.3% and 88.4%, and F1-scores of 91.0% and 88.0%, respectively. The corresponding average values for manual inspection were 89.9%, 89.8%, and 89.5%, respectively. In comparison, YOLOv11n-DAL achieved a precision of 95.2%, recall of 94.5%, and F1-score of 94.8%, representing improvements of 5.3, 4.7, and 5.3 percentage points, respectively, over the average manual inspection performance.

These results indicate that YOLOv11n-DAL provides superior overall detection performance compared with manual inspection under conveyor-belt operating conditions. Moreover, the differences between the two inspectors indicate a certain degree of variability in manual inspection, which may be attributed to subjective judgment and differences in visual perception. In contrast, YOLOv11n-DAL provides a more consistent and automated detection process. Therefore, the proposed model demonstrates good potential for automated egg defect detection in conveyor-belt inspection scenarios.

Training Stability Analysis

To evaluate the robustness of YOLOv11n-DAL against training randomness, both YOLOv11n and YOLOv11n-DAL were independently trained five times using different random seeds (0–4), while maintaining the same dataset split, training strategy, and hyperparameter settings. The mean ± standard deviation of the five runs, together with paired t-test results and effect sizes, are summarized in Table 7.

Table 7

Training stability comparison between YOLOv11n and YOLOv11n-DAL over five runs

Metric	YOLOv11n (Mean ± SD)	YOLOv11n-DAL (Mean ± SD)	Δ	p-value	Cohen's d_z
P (%)	94.4 ± 0.7	94.9 ± 0.7	+0.5	0.373	0.5
R (%)	93.3 ± 0.2	94.1 ± 1.0	+0.8	0.187	0.7
F1 (%)	93.8 ± 0.4	94.5 ± 0.4	+0.7	0.047	1.3
mAP50 (%)	96.0 ± 0.2	96.8 ± 0.2	+0.7	0.010	2.0
mAP50–95 (%)	94.9 ± 0.2	95.6 ± 0.4	+0.7	0.028	1.5

Note: Values are presented as mean ± standard deviation over five independent runs with different random seeds (0–4).
Δ represents the performance improvement of YOLOv11n-DAL compared with YOLOv11n.
Statistical significance was evaluated using paired t-tests.
Cohen's dz represents the paired effect size.

As shown in Table 7, YOLOv11n-DAL achieved higher average performance than YOLOv11n across all evaluation metrics. Specifically, the proposed model improved Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95 by 0.5, 0.8, 0.7, 0.7, and 0.7 percentage points, respectively. The relatively small standard deviations (<1.1%) across repeated experiments indicate that both models exhibited stable training behavior.

The paired t-test results showed that the improvements in F1-score, mAP@0.5, and mAP@0.5:0.95 achieved statistical significance in the five repeated runs, with relatively large effect sizes. Although Precision and Recall improvements did not reach statistical significance, they remained consistently higher than the baseline model. Overall, the repeated experiments demonstrate that the performance gains of YOLOv11n-DAL are reproducible and not caused by a single random training process.

DISCUSSION

The experiments show that YOLOv11n-DAL improves detection accuracy while reducing model complexity rather than relying on a larger network. ADown contributes mainly to computational reduction, C3k2_DCNv4 strengthens the representation of irregular shell defects, and Detect_LSDECD reduces redundant prediction operations while enhancing fine-grained feature extraction. Their complementary effects explain the simultaneous improvements in F1-score and mAP and the reductions in parameters, GFLOPs, and model size.

It is noteworthy that the difference between mAP@0.5 and mAP@0.5:0.95 is relatively small in this study. For YOLOv11n-DAL, the gap between the two metrics was only 0.9 percentage points (96.7% vs. 95.8%), indicating that the model maintained stable localization performance under stricter IoU thresholds. This phenomenon is mainly attributed to the characteristics of the proposed dataset and annotation strategy. Unlike defect-region detection tasks that require precise localization of irregular crack or stain areas, the bounding boxes in this study were annotated for complete egg regions. Since eggs exhibit relatively regular elliptical shapes and clear boundaries from the background, the localization difficulty is relatively low, allowing the model to maintain high IoU overlap even under stricter evaluation criteria. In addition, the C3k2_DCNv4 module improves structural feature modeling through adaptive sampling, which further contributes to accurate object localization.

The dirty category remained more difficult than damaged and good because slight stains, natural eggshell spots, shadows, and illumination-induced dark regions may share similar visual characteristics. Partial occlusion and single-view acquisition also limit the visibility of some defects. Moreover, the dataset was collected using one device within a limited region and period, which may affect the generalization ability under different acquisition conditions. Therefore, further validation using multi-source datasets and diverse imaging conditions is required before large-scale deployment. In addition, the statistical analysis was conducted based on five repeated training runs (n=5), and the limited sample size may reduce the statistical power of significance tests. Moreover, no correction for multiple comparisons was applied; therefore, the statistical results should be interpreted as supportive evidence rather than definitive proof of superiority.

CONCLUSIONS

This study proposed YOLOv11n-DAL, a lightweight detection framework for egg external-defect inspection under conveyor-belt conditions. By integrating C3k2_DCNv4, ADown, and Detect_LSDECD into YOLOv11n, the proposed model enhanced irregular defect feature representation, improved feature preservation during downsampling, and reduced redundant prediction computation.

Compared with the baseline YOLOv11n, YOLOv11n-DAL achieved improved detection accuracy while reducing model complexity, with the parameters, GFLOPs, and model size reduced from 2.6 M, 6.3 GFLOPs, and 5.2 MB to 1.7 M, 4.8 GFLOPs, and 3.8 MB, respectively. The ablation experiments, comparative evaluations, and stability analysis demonstrated the effectiveness and reproducibility of the proposed optimization strategy.

The proposed model provides a feasible lightweight solution for automated egg external-defect detection in resource-constrained conveyor-belt sorting scenarios. Future work will focus on validating the model under more diverse acquisition conditions and conducting deployment tests on practical edge devices.

REFERENCES

[1] Atwa, E. M., Xu, S., Rashwan, A. K., Abdelshafy, A. M., ElMasry, G., Al-Rejaie, S., Xu, H., Lin, H., Pan, J., (2024). Advances in emerging non-destructive technologies for detecting raw egg freshness: a comprehensive review, *Foods*, vol.13, no.22, 3563, Switzerland. https://doi.org/10.3390/foods13223563

[2] Botta, B., Gattam, S. S. R., Datta, A. K., (2022). Eggshell crack detection using deep convolutional neural networks, *Journal of Food Engineering*, vol.315, 110798, United Kingdom. https://doi.org/10.1016/j.jfoodeng.2021.110798

[3] Bing, Y., Yu, X., Lin, Z., Yu, F., (2025). PAME-YOLO: A model for apple leaf lesion detection in complex environments based on improved YOLOv8s, *INMATEH - Agricultural Engineering*, vol.75, no.1, pp.1179–1192, Romania. https://doi.org/10.35633/inmateh-75-97

[4] Chen, S. Y., Hsu, S. H., Ko, C. Y., Hsu, K. H., (2023). Real-time defect and freshness inspection on chicken eggs using hyperspectral imaging, *Food Control*, vol.150, 109716, United Kingdom. https://doi.org/10.1016/j.foodcont.2023.109716

[5] Çengel, T. A., Gencturk, B., Yasin, E. T., Yildiz, M. B., Cinar, I., Koklu, M., (2025). Automating egg damage detection for improved quality control in the food industry using deep learning, *Journal of Food Science*, vol.90, no.1, e17553, United States. https://doi.org/10.1111/1750-3841.17553

[6] Chen, Z., He, Z., Lu, Z. M., (2024). DEA-Net: single image dehazing based on detail-enhanced convolution and content-guided attention, *IEEE Transactions on Image Processing*, vol.33, pp.1002–1015, United States. https://doi.org/10.1109/TIP.2024.3354108

[7] Dai, J., Qi, H., Xiong, Y., Li, Y., Zhang, G., Hu, H., Wei, Y., (2017). Deformable convolutional networks, Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp.764–773, Italy. https://doi.org/10.1109/ICCV.2017.89

[8] Gautron, J., Stapane, L., Le Roy, N., Nys, Y., Rodriguez-Navarro, A. B., Hincke, M. T., (2021). Avian eggshell biomineralization: an update on its structure, mineralogy and protein tool kit, *BMC Molecular and Cell Biology*, vol.22, 11, United Kingdom. https://doi.org/10.1186/s12860-021-00350-0

[9] Gupta, C., Gill, N. S., Gulia, P., Duhan, S., Alduaiji, N., Shukla, P. K., Dwivedi, A., (2025). Deep vision in agriculture: assessing the function of YOLO in the classification of plant leaf diseases, *BioData Mining*, vol.18, 91, United Kingdom. https://doi.org/10.1186/s13040-025-00497-y

[10] Huang, Y., Luo, Y., Cao, Y., Lin, X., Wei, H., Wu, M., Yang, X., Zhao, Z., (2023). Damage detection of unwashed eggs through video and deep learning, *Foods*, vol.12, no.11, 2179, Switzerland. https://doi.org/10.3390/foods12112179

[11] Jocher, G., Chaurasia, A., Qiu, J., (2024). Ultralytics YOLO11, *Ultralytics software documentation*, United States. Available at: https://docs.ultralytics.com/models/yolo11/

[12] Kanna, S. K., Ramalingam, K., Pazhanivelan, P., Jagadeeswaran, R., Prabu, P. C., (2024). YOLO deep learning algorithm for object detection in agriculture: a review, *Journal of Agricultural Engineering*, vol.55, no.4, 1641, Italy. https://doi.org/10.4081/jae.2024.1641

[13] Ma, F., Li, S., Tan, J., Li, Y., (2025). YOLO-TRS: An improved YOLO11 for tomato fruit ripeness and stem detection, *INMATEH - Agricultural Engineering*, vol.77, no.3, pp.1131–1144, Romania. https://doi.org/10.35633/inmateh-77-91

[14] Shi, W., Zhang, D., Wei, N., Ge, Y., Chen, J., Huang, T., Chen, Y., Zhang, S., Hong, B., Yuan, W., Zhang, B., (2025). An improved YOLOv11n-based GENET model for vision-based missing-seed detection and counting in an oblique hook-shaped spoon-type small precision seed metering device, *INMATEH - Agricultural Engineering*, vol.77, no.3, pp.490–501, Romania. https://doi.org/10.35633/inmateh-77-40

[15] So, J. H., Kim, H. S., (2022). Current advances in detection of abnormal egg: a review, *Journal of Animal Science and Technology*, vol.64, no.5, pp.813–829, Korea. https://doi.org/10.5187/jast.2022.e52

[16] Tang, F., Zhang, Y., Zhang, Y., Yang, L., (2026). An improved RT-DETR for stable and real-time defective egg detection in edge computing environments, *INMATEH - Agricultural Engineering*, vol.78, no.1, pp.644–655, Romania. https://doi.org/10.35633/inmateh-78-52

[17] Turkoglu, M., (2021). Defective egg detection based on deep features and Bidirectional Long-Short-Term-Memory, *Computers and Electronics in Agriculture*, vol.185, 106152, Netherlands. https://doi.org/10.1016/j.compag.2021.106152

[18] Wu, Z., Zhang, H., Fang, C., (2025). Research on machine vision online monitoring system for egg production and quality in cage environment, *Poultry Science*, vol.104, 104552, United States. https://doi.org/10.1016/j.psj.2024.104552

[19] Wang, C. Y., Yeh, I. H., Liao, H. Y. M., (2024). YOLOv9: learning what you want to learn using programmable gradient information, Proceedings of the European Conference on Computer Vision (ECCV), pp.1–21, Italy. https://doi.org/10.1007/978-3-031-72751-1_1

[20] Wang, P., Li, C., Wei, L., Li, Y., Guo, L., (2026). Improved YOLOv11 for lightweight and accurate detection of flax diseases and pests in complex field environments, *Agriculture*, vol.16, no.7, 741, Switzerland. https://doi.org/10.3390/agriculture16070741

[21] Xiong, Y., Li, Z., Chen, Y., Wang, F., Zhu, X., Luo, J., Wang, W., Lu, T., Li, H., Qiao, Y., Lu, L., Zhou, J., Dai, J., (2024). Efficient deformable convnets: Rethinking dynamic and sparse operator for vision applications, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.5652–5661. https://doi.org/10.1109/CVPR52733.2024.00540.

[22] Yang, X., Bist, R. B., Subedi, S., Chai, L., (2023). A computer vision-based automatic system for egg grading and defect detection, *Animals*, vol.13, no.14, 2354, Switzerland. https://doi.org/10.3390/ani13142354

[23] Yin, J. J., Kang, J. Q., Xiao, D. Q., (2024). Lightweight detection algorithm for duck egg cracks based on improved YOLOv5l, *Transactions of the Chinese Society of Agricultural Engineering*, vol.40, no.5, pp.216–223, China. https://doi.org/10.11975/j.issn.1002-6819.202307259

[24] Yu, P., Shen, C., Cheng, J., Yin, X., Liu, C., Yu, Z., (2025). Advances in nondestructive technologies for external eggshell quality evaluation, *Sensors*, vol.25, no.18, 5796, Switzerland. https://doi.org/10.3390/s25185796

[25] Zhao, Z. X., Wei, H. F., Huang, Y., Huang, X. P., Mi, Y. L., Luo, Y. F., (2023). Online real-time detection system for broken shell eggs based on improved YOLOv7, *Transactions of the Chinese Society of Agricultural Engineering*, vol.39, no.20, pp.255–265, China. https://doi.org/10.11975/j.issn.1002-6819.202305228

[26] Zhu, G., Qi, H., Lv, K., (2025). DGYOLOv8: an enhanced model for steel surface defect detection based on YOLOv8, *Mathematics*, vol.13, no.5, 831, Switzerland. https://doi.org/10.3390/math13050831