

OPTIMIZATION OF DAMAGED CORN KERNEL RECOGNITION ALGORITHM BASED ON A DUAL-LIGHT SYSTEM

基于双光系统的破损玉米籽粒识别算法优化

Yuchao REN¹⁾, Bin QI ^{*1,2)}, Xiaoming SUN^{*1,2)}, Jiyuan SUN¹⁾, Tengyun MA¹⁾, Yuanqi LIU¹⁾, Bohan ZHANG¹⁾, Qiong WU³⁾

¹⁾ College of Agricultural Engineering and Food Science, Shandong University of Technology/ China;

²⁾ School of Fine Arts and Design, Shandong University of Technology, Zibo City, China.

³⁾ Zibo Normal College/ China

Tel: +86-15966964198; E-mail: sdutid@163.com

Tel: +86-13581019376; E-mail: 3282110@qq.com

DOI: <https://doi.org/10.35633/inmateh-75-67>

Keywords: Corn Kernel, Object Detection, Light Angle, YOLOv8n

ABSTRACT

To enhance real-time detection of corn breakage rate under dim conditions, this study designed a dual-light (top/backlight) sampling system. By comparing four datasets (top-scattered, top-clustered, backlight-scattered, backlight-clustered), the algorithm optimized with backlight-scattered data achieved optimal accuracy (79.6%). A lightweight YOLOv8n_gcd model was proposed, integrating Ghost convolution in the backbone to reduce redundancy, attention mechanisms for feature enhancement, and depthwise separable convolutions in the neck. The optimized model reduced FLOPs by 24% and increased FPS by 165%, offering an efficient, low-cost solution for agricultural quality inspection with theoretical and practical value.

摘要

为了增强在昏暗条件下玉米破损率的实时检测，本研究设计了一种双光（顶部/背光）采样系统。通过比较四个数据集（顶光-籽粒分散、顶光-籽粒聚集、背光-籽粒分散、背光籽粒聚集），得出背光籽粒分散的数据优化的算法达到了最佳（79.6%）之后利用该数据集训练出一种轻量级的 YOLOv8n_gcd 模型，将 Ghost 卷积集成在骨干网中以减少冗余，将注意力机制用于特征增强，并在颈部进行深度可分离卷积。优化后的模型将 FLOP 降低了 24%，FPS 提高了 165%，为农业质量检测提供了一种高效、低成本的解决方案，具有理论和实践价值。

INTRODUCTION

Corn holds the position as the largest - scale grain crop in China (Cui *et al.*, 2019). With the progressive development of agricultural mechanization, the mode of corn harvesting in China is undergoing a significant transition from ear - based harvesting to direct kernel – harvesting (Li, 2017; Zhao *et al.*, 2020). This shift in the harvesting method represents a crucial step in modernizing corn production, enabling a more efficient and cost-effective approach. It has thus become an inevitable trend in the mechanized production of corn in the country (Zhu *et al.*, 2021). The operational quality of corn harvesting exerts a profound impact on both its yield and economic value. The breakage rate and impurity content serve as pivotal metrics for evaluating the performance of direct - kernel harvesters (Zhao *et al.*, 2025; Cui Y.S., 2024; Yang *et al.*, 2018). However, the traditional manual detection method is fraught with limitations. It is characterized by high labor intensity and low detection efficiency, which not only burdens the workforce but also fails to provide real-time and accurate feedback of the breakage rate and impurity content to the automatic control system of the harvester. As a result, timely adjustments to the operation parameters cannot be made, leading to potential substantial harvesting losses (Xu *et al.*, 2021; Wu *et al.*, 2024). This in turn has significantly impeded the advancement of intelligent harvesting technologies for domestic corn direct - kernel harvesters.

Convolutional Neural Networks (CNvolutional Neural Networks, CNNs) in deep learning have demonstrated exceptional performance improvements in cutting-edge fields such as object recognition, image classification, and image segmentation.

Yuchao Ren, M.S. Agr.; Bin Qi Assoc. Prof. Ph.D. Eng.; Xiaoming Sun, M.S. Eng., Lecturer; Jiyuan Sun, M.S. Agr.; Tengyun Ma, M.S. Agr.; Yuanqi Liu, M.S. Agr.; Bohan Zhang, M.S. Agr.; Qiong Wu, Senior Art and Craft Specialist.

Liu *et al.* showed significant advantages in training multi-scale image input models compared to single-scale scene classification strategies (Liu *et al.*, 2018). Since the introduction of ResNet, residual networks have successfully addressed the long-standing vanishing gradient problem enabling unrestricted network depth and garnering widespread attention in computer vision (He *et al.*, 2016). However, deep learning-based image classification tasks often suffer from excessive parameters and bulky models, posing challenges for deployment on mobile devices. To address this, Andrew *et al.* proposed MobileNets, a lightweight convolutional neural network that achieves efficient computation on mobile and embedded systems while reducing memory consumption, thereby streamlining model parameters and computational load (Howard *et al.*, 2017). However, this simplification comes at the cost of some accuracy loss. In MobileNetV2, Mark *et al.* introduced an inverted residual structure, achieving higher accuracy with the same computational cost while minimizing information loss (Sandler *et al.*, 2018).

In the field of damaged corn kernel identification, Velesaca *et al.* proposed a two-stage static detection algorithm for corn kernel breakage (Velesaca *et al.*, 2020). This method first segments individual kernels from clustered corn images using Mask R-CNN, then classifies them via a custom-designed network (CK-CNN). Han *et al.* introduced an independent component analysis (ICA)-based method for germ feature detection, selecting germ-representative components from RGB color space and integrating nine additional germ-area features, yielding a minimal area error of 0.7% compared to manual inspection (Han *et al.*, 2010). Li *et al.* developed an image acquisition device and an improved YOLOv4-tiny model for broken kernel detection, achieving 93.5% and 93% precision for intact and damaged kernels, respectively, with lower detection error than manual methods, demonstrating real-time applicability (Li *et al.*, 2021). Xu *et al.* selected corn varieties Denghai 518, Xundan 20, and Zhengdan 958, constructing a CNN model using the Keras deep learning framework, achieving an average recognition rate of 95.49% (Xu *et al.*, 2020). Quan *et al.* investigated corn kernel selection and classification, developing a lightweight convolutional neural network. By optimizing prototype parameters, they achieved optimal performance with a detection accuracy of 96.50% and an effective sorting rate of 97.51% for four categories of kernels (high-quality, rejected, germ-side, and endosperm-side) (Quan *et al.*, 2020).

As mentioned earlier, it is necessary to research an algorithm that is adaptable to dark conditions and is lighter and easier to deploy on mobile devices.

DUAL LIGHT DEVICE DESIGN

Lighting System Design

To address image acquisition challenges in low-light environments, this study designed a dual-light system with engineered multi-angle coordination (upper and lower illumination). Comparative experiments were conducted using datasets under distinct lighting conditions to evaluate optimal angles for image capture. Hybrid-angle lighting (simultaneous upper/lower illumination) was excluded to ensure experimental comparability. The core light source is the Leichu Lighting 600×600 direct-emitting LED panel which named Model LZ-600D, whose technical specifications are compared with industry standards in Table 1.

Table 1

Comparison of Main Light Source Technical Parameters and Industry Standards

Parameter	Light emission type	Rated power [W]	Luminous flux [lm]	Illuminance uniformity [%]	Color temperature [K]	Color Rendering Index (CRI)
Leichu 600×600 LED Panel	COB-integrated direct/edge-lit	80	9,600	92 (at 0.5 m distance)	6,500 ± 200	Ra ≥ 96
Industry Standard (Industrial Grade)	Direct/edge-lit	50–100	8,000–12,000	≥85	5,000–6,500 (recommended)	Ra ≥ 90 (for precision inspection)

Mechanical Structure and Optical Layout

A modular dark chamber (80 cm × 50 cm × 50 cm) was constructed to simulate post-threshing conditions (ambient light <10 lux). The chamber's interior was coated with black light-absorbing paint (reflectivity <2%), and labyrinthine light seals (attenuation >99%) were installed at joints to eliminate stray light interference. A circular aperture (8 cm diameter) was integrated at the top for imaging.

The lighting system comprised two configurations (Fig. 2):

Top Lighting: A ring-shaped array of 24 COB LED modules (3W each) with a three-layer homogenizing structure (diffuser plate, micro-prism film, anti-glare mesh) to reduce illuminance gradients to <5%.

Backlighting: An identical LED array and homogenizer positioned at the chamber base to create a planar light source.

A co-axial optical path (light source–camera–sample alignment) was implemented to minimize edge shadows and ensure accurate kernel morphology recognition.

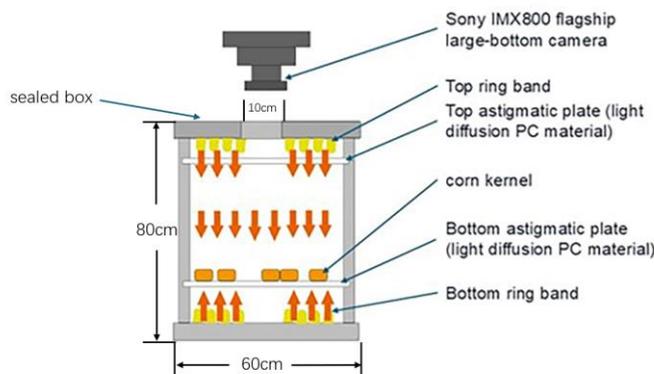


Fig. 2 – Schematic layout of the sampling device

Imaging System Design

The imaging system featured:

1) Camera: A Sony IMX800 image sensor (1/1.49", 54 MP) with 82% quantum efficiency at 550 nm, readout noise of $2.3e^-$, and dynamic range of 73 dB.

2) Lens: A Computar MLH-10X fixed-focus lens (25 mm focal length, f/2.8 aperture).

3) Stage: A three-axis precision stage with ± 50 mm horizontal range, <0.01 mm repeatability, and vertical electric focusing (0–100 mm travel). A laser rangefinder (0.1 mm accuracy) and dual-axis gimbal (15° pitch, 10° roll correction) were integrated to mitigate mechanical misalignment and vibration-induced imaging errors.

MATERIALS AND METHODS

Target Materials

The corn kernels harvested in the Northeast region of China during late September 2024 were chosen as the representative samples for this study. Two distinct types of corn kernels were the focus of detection: intact corn kernels, as illustrated in Fig. 3 (a), and damaged corn kernels, as shown in Fig. 3 (b).



Fig. 3 – Types of corn kernels

a) Complete corn kernels; b) Damaged corn kernels

During the corn - threshing process, kernel damage can occur due to multiple factors. Mechanically, the use of certain threshing machines, such as those equipped with round - headed spike - tooth mechanisms, can cause physical damage to the kernels. Additionally, improper operational practices, including sub-optimal moisture content levels (either too high or too low) and uneven feeding of corn materials into the threshing machine, can also contribute to kernel breakage.

Image Acquisition

Image collection was conducted within the established image - acquisition system. Initially, the enclosure of the system was opened, and the corn kernels designated for imaging were carefully placed inside. Subsequently, the enclosure was sealed to create a controlled, relatively dark environment that mimics the real-world conditions of post-threshing. Depending on the experimental setup, either the top - light or the back - light was activated. The height of the camera was then adjusted until the entire area of the corn kernels was within the field of view. For each batch of corn kernels, two images were captured: one with the top - light illumination and another with the back - light illumination. After image acquisition, the corn kernels were replaced, and the process was repeated. In total, 684 image pairs were collected.

Typically, the quality of the dataset significantly influences the predictive performance of subsequent deep - learning algorithms. To enhance the generalization and robustness of the network model, conventional data - augmentation techniques, such as noise injection, exposure adjustment, horizontal and vertical flipping, and random rotation, are often employed. However, in this experiment, the primary objective was to compare the efficacy of top - light and back - light illumination for detecting damaged corn kernels. Therefore, data - augmentation techniques were not applied to maintain the integrity of the original lighting - condition data.

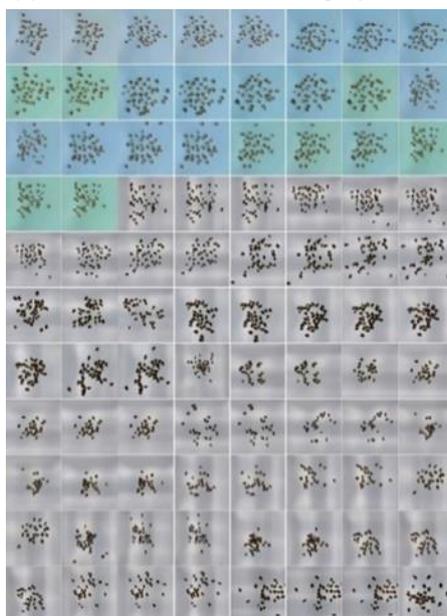


Fig. 4 – Sample images collected during the experiment

Image Calibration and Processing

After image acquisition, the datasets were categorized into four classes: top-scattered, top-clustered, backlight-scattered, and backlight-clustered. Subsequently, images within each subfolder were partitioned into training, validation, and test sets at a ratio of 6:2:2 to ensure balanced data distribution.

The images in each set were then imported into the Labelling software, a widely - used tool in the field of image recognition for manual annotation. In this study, the only objects of interest were intact and damaged corn kernels. Therefore, only these two types of kernels were annotated within the images, enabling the training of the subsequent detection algorithm.

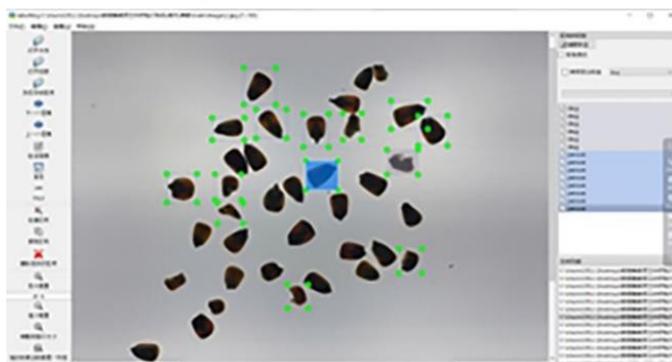


Fig. 5 – Example of an image being annotated in the Labelling software

ALGORITHM OPTIMIZATION AND TRAINING

YOLOv8 Model

Given the real-world application scenario of detecting damaged corn kernels, the YOLOv8n algorithm was selected for this study.

Renowned for its high - speed and efficient object - detection capabilities, YOLOv8n has found extensive applications across various domains. It is a component of the YOLOv8 network model, which supports multiple computer - vision tasks, including image classification, object detection, and instance segmentation. The YOLOv8 architecture consists of five main components: Input, Backbone, Neck, Head, and Output.

Compared to the full - scale YOLOv8 model, YOLOv8n has a more streamlined architecture with fewer model parameters and reduced computational requirements. This makes it particularly suitable for deployment in resource - constrained environments while still maintaining a relatively high detection speed. Despite its lightweight design, YOLOv8n can achieve competitive accuracy in object - detection tasks.

GhostNet

The backbone network architecture of the YOLOv8 model utilizes CSPDarknet. The significant depth and width of this backbone network, while suitable for high precision, incur high computational costs that degrade YOLOv8's overall detection speed. To address this limitation, this chapter introduces the lightweight network GhostNet, replacing the CSPDarknet in YOLOv8 with the hierarchical structure of GhostNet.

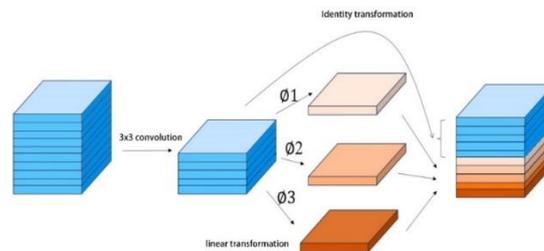


Fig. 6 – Ghost Convolution Process

CBAM Attention Mechanism

When detecting damaged corn kernels, the input images often contain not only the target objects but also substantial complex background information. The deep architecture of YOLOv8, after undergoing multiple convolutional layers, tends to neglect critical target feature information while disproportionately focusing on background features. To address this issue, this chapter integrates a spatial-channel combined attention mechanism—Convolutional Block Attention Module (CBAM) (*Woo et al., 2018*)—into the backbone network of YOLOv8. This mechanism aims to enhance the network's ability to precisely localize regions of interest (ROIs) and suppress the influence of background noise in remote sensing images. The core principle of the attention mechanism lies in dynamically amplifying the saliency of specific spatial and channel-wise features within input images, thereby refining the extraction of key discriminative features for improved detection accuracy.

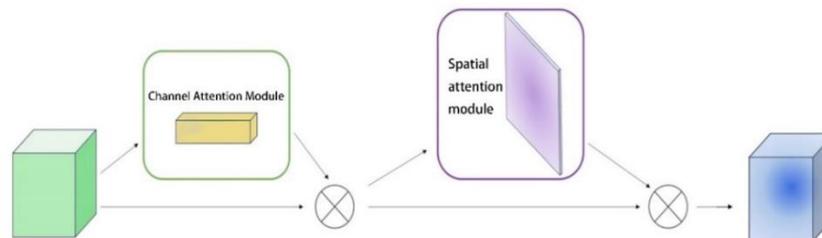


Fig. 7 – CBAM Attention Mechanism

Depthwise Separable Convolution

To further enhance the detection accuracy of the model, an attention mechanism can be introduced in the backbone of the original network. However, the attention mechanism may negatively impact inference speed. Therefore, Depthwise Separable Convolution (DWConv) is incorporated into the neck for localized optimization (*Woo et al., 2018*).

In the modified YOLOv8 model, standard convolutions in the backbone network are replaced with DWConv. The computation is divided into two steps: depthwise convolution (applying a single filter per input channel) followed by pointwise convolution (combining outputs across channels via 1×1 convolutions). Compared to standard convolutions, this approach drastically reduces computational complexity and accelerates model inference speed. The structure of this modification is illustrated in Fig. 8.

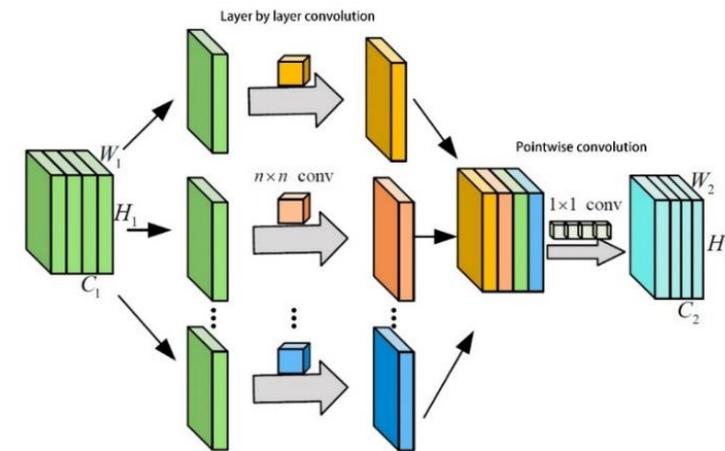


Fig. 8 – Depth Separable Convolutional Structure

MODEL TRAINING AND RESULT ANALYSIS

Training Environment

The experimental setup was configured with a 64-bit Windows 10 operating system. The hardware specifications included a 12th-generation Intel(R) Core (TM) i7-12700KF CPU operating at a base frequency of 3.61 GHz, 32 GB of RAM, and an NVIDIA 3080Ti graphics processing unit (GPU) with 12 GB of dedicated video memory. The deep-learning framework employed was Python 3.11.0, in conjunction with Cuda 12.4 for GPU-accelerated computations. Python 3.8.0 was used as the programming language, and the development environment was PyCharm Community Edition.

Evaluation Indicators

In the process of model training, two key performance indicators, precision and recall, play a vital role in assessing the algorithm's effectiveness. Precision measures the proportion of correctly identified positive samples among all samples predicted as positive, while recall represents the proportion of correctly identified positive samples among all actual positive samples. The mathematical formulas for calculating precision (P) and recall (R) are as follows:

$$P = \frac{TP}{TP+FP} \quad (1)$$

$$R = \frac{TP}{TP+FN} \quad (2)$$

where: *TP* (True Positive) refers to positive samples that are correctly classified as positive; *TN* (True Negative) represents negative samples that are correctly classified as negative; *FP* (False Positive) denotes negative samples that are incorrectly classified as positive; and *FN* (False Negative) indicates positive samples that are incorrectly classified as negative.

Mean Average Precision (*MAP*) is another crucial metric for evaluating the overall performance of the network model. It provides a comprehensive measure of the model's accuracy across different confidence thresholds. The formula for calculating *MAP* is as follows:

$$AP = \int_0^1 p(r) dr \quad (3)$$

$$MAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (4)$$

where: *AP* (Average Precision) represents the average precision for a specific class of objects; and *k* is the total number of classes in the dataset.

In this experiment, four evaluation metrics were utilized to analyze the experimental results and compare them with the baseline network performance. These metrics include MAP_{b50 - 95}, MAP_{m50 - 95}, computational complexity, and inference time. The notation 50 - 95 represents the average MAP calculated at different Intersection over Union (IoU) thresholds, ranging from 50% to 95% with a step size of 5%.

Comparative experiment of datasets from different perspectives

To determine which lighting angle, top - light or back - light, is more suitable for detecting damaged corn kernels in real-world scenarios, the YOLOv8n algorithm was trained using four different datasets: top-scattered, backlight-scattered, top-clustered, and backlight-clustered. All training was conducted under identical experimental conditions, using the hyperparameters tuned from the original model and maintaining consistent training epochs and learning rates. The experimental verification results are presented in Table 2.

Table 2

Verification results of algorithms trained with different datasets

datasets	Precision/%	Recall/%	mAP _{0.5~0.95} (%)	FPS/F-s-1
<i>Top-scattered</i>	89.8	67	70.7	35
<i>Top-clustered</i>	76.6	77.7	61.4	39
<i>Backlight-scattered</i>	85.7	70.6	79.6	31
<i>Backlight-clustered</i>	80.6	74.1	75.4	30

As indicated by the data in Table 2, the Dispersal in the Backlight dataset performs the best, along with the fastest recognition speed. Through horizontal comparison, it was evident that discrete-kernel datasets were more conducive to training, resulting in more accurate models. Moreover, back-light-illuminated datasets generally outperformed top-light-illuminated datasets in terms of training efficiency and model accuracy. Therefore, backlight-scattered dataset was used to train the improved algorithm model.

Data Enhancement

After identifying the backlight-scattered dataset as the optimal training dataset, data preprocessing was performed to enhance sample diversity and enable the network to learn features from multiple perspectives before feeding it into the training pipeline. This process improves the model's ability to analyze feature data and strengthens its generalization capability. Data augmentation techniques—such as rotation, flipping, blurring, and brightness adjustment—were performed on the dataset samples, as illustrated in Fig. 9.

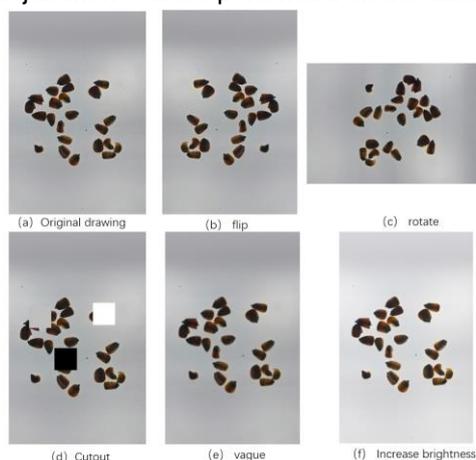


Fig. 9 – Data enhanced image

RESULTS

Ablation Experiment

To verify the effectiveness of the improvements proposed in this article, ablation experiments were conducted to validate the effectiveness of each part. The experimental results are shown in Table 3, where YOLOv8n_g represents the use of Ghost convolution on the basis of YOLOv8n, YOLOv8n_gc represents the use of attention mechanism on the basis of YOLOv8n_g, and YOLOv8n_gcd represents the addition of depthwise separable convolution on the basis of YOLOv8n_gc. Compared to YOLOv8n GFLOPS, YOLOv8n_g has decreased from 8.1 to 5.9, the model's accuracy mAP has decreased from 86.6% to 84.6%, and FPS has increased from 43 to 68. Although using Ghost convolution can significantly improve the model's detection efficiency, it will result in a loss of accuracy.

Compared to YOLOv8n_g, YOLOv8n_gc showed a 0.9 percentage point increase in accuracy and an 8% decrease in FPS with a 0.2 increase in GFLOPs. The attention mechanism can enhance the model's feature extraction ability for targets and improve detection accuracy. Compared to YOLOv8n_gcd, YOLOv8n_gcd reduces computational complexity by 7.97%, increases detection speed by 11fps, and reduces mAP value by 1.05%. Compared to YOLOv8n, YOLOv8n_gcw significantly improves FPS by 60% while slightly reducing model accuracy.

Table 3

Results of ablation experiment

Model	Ghost	CBAM	DWConv	GFLOPs	FPS	mAP
YOLOv8n	-	-	-	8.1	43	86.6%
YOLOv8n_g	√	-	-	5.9	68	84.6%
YOLOv8n_gc	√	√	-	6.1	60	85.5%
YOLOv8n_gcd	√	√	√	6.1	71	84.95%

The specific training dynamics of the YOLOv8n_gcw model are illustrated in Fig. 10. By observing the mAP@50 and mAP@50-95 metrics in Figure 10, it is evident that the model's average precision values stabilize around the 200th epoch, with no significant improvement thereafter. This indicates that the model has converged to its optimal state and does not benefit from further training iterations.

Concurrently, the evaluation metrics for Precision and Recall also exhibit stable and unchanging trends at this stage, suggesting that the model has achieved an optimal balance between detecting true positives and minimizing false positives/negatives. Therefore, additional training epochs are unnecessary, as they would not yield meaningful performance gains.

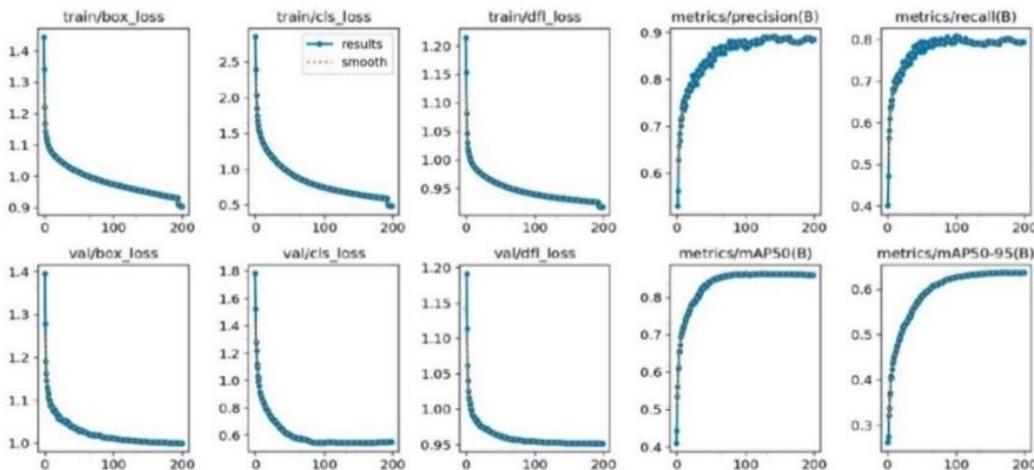


Fig. 10 – YOLOv8n_gcd Training Process Diagram.

Contrast Test

To further demonstrate the effectiveness of the algorithm presented in this chapter, comparative experiments were conducted against classical deep learning object detection algorithms, including Faster R-CNN, SSD, and YOLOv7 (Wang et al., 2023). The experimental results are shown in Table 4.

Table 4

Comparative experiment of different object detection algorithms

Model	GFLOPs	mAP@0.5	Precision	Recall
Faster-RCNN	227.8	80.6%	81.8%	74.6%
SSD	63.14	78.6%	79.1%	70.9%
YOLOv7	105.4	81.5%	82.4%	73.5%
YOLOv8n_gcd	6.1	84.95	84.8%	76.2%

The experimental results in Table 4 clearly demonstrate that the optimized model proposed in this study achieves significantly superior performance in reducing floating-point operations (FLOPs), with its FLOPs far lower than those of the other three benchmark models. In other aspects, compared to Faster R-CNN, SSD, and YOLOv7, the optimized model exhibits improvements in precision by 3%, 5.7%, and 2.4%, respectively; recall by 1.6%, 5.3%, and 2.7%, respectively; and mean average precision (mAP@0.5) by 4.35%, 6.35%, and 3.45%, respectively. Notably, the YOLOv8n_gcd model maintains a compact architecture with reduced computational demands, enabling rapid deployment on resource-constrained platforms. These enhancements collectively validate the superior performance and broad applicability of the proposed improved model in real-world applications.

CONCLUSIONS AND DISCUSSIONS

This study designed a dual-light device to collect datasets, generating four distinct datasets: top-lighting-dispersed, top-lighting-aggregated, backlighting-dispersed, and backlighting-aggregated, which were used to train the same algorithm. The results revealed that the backlighting-dispersed dataset achieved the highest accuracy and was most suitable for training. Building on this, an optimized YOLOv8-based model, termed YOLOv8n_gcd, was proposed. This model integrates Ghost convolutions into the backbone network to minimize redundant computations, incorporates attention mechanisms to enhance feature extraction for damaged regions, and replaces standard convolutions with depthwise separable convolutions in the Neck section. These localized optimizations achieve a balance between accuracy and lightweight efficiency. Comparative experiments demonstrated that the YOLOv8n_gcd model outperforms Faster R-CNN, SSD, and YOLOv7, with precision improvements of 3%, 5.7%, and 2.4%, recall improvements of 1.6%, 5.3%, and 2.7%, and mAP@0.5 gains of 4.35%, 6.35%, and 3.45%, respectively. Notably, the optimized YOLOv8s model retains a compact architecture with significantly reduced computational demands, enabling rapid deployment on resource-limited platforms. These advancements in precision, speed, and robustness validate the superiority and broad applicability of the proposed model in practical agricultural scenarios, offering a high-performance solution for real-time quality inspection tasks while maintaining adaptability to industrial constraints.

ACKNOWLEDGEMENT

This work was supported by the Shandong Province Science and Technology Minor Enterprises Innovation Ability Promotion Project (Project number: 2023TSGC1014), General Program of China Higher Education Association (23NL0406) and the Key Project of Undergraduate Teaching Reform Research in Shandong Province (Z2022263). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

- [1] Cui, T., Fan, C. L., Zhang, D. X., Yang, L., Li, Y. B., & Zhao, H. H. (2019). Research progress on mechanical harvesting technology of maize (玉米机械化收获技术研究进展分析). *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 50(12), pp. 1–13, China.
- [2] Cui, Y. S. (2024). Technological progress of domestic corn combine harvesters in China (国内玉米联合收获机技术进展). *Agricultural Engineering*, Vol. 14(5), pp. 22-26, China.
- [3] Han, Z. Z., Zhao, Y. G., & Yang, J. Z. (2010). Corn embryo feature detection based on independent components of RGB kernel images (基于籽粒 RGB 图像独立分量的玉米胚部特征检测). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 26(03), 222–226+389, China.
- [4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp. 770–778. China.
- [5] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv*. <https://doi.org/10.48550/arXiv.1704.04861>, China.
- [6] Quan, L. Z., Wang, J. Y., Wang, Q., Xiao, Y. H., & Feng, H. Q. (2020). Corn quality selection device based on electromagnetic vibration and convolutional neural network (基于电磁振动与卷积神经网络的玉米品质精选装置). *Journal of Jiangsu University (Natural Science Edition)*, Vol. 41(03), pp. 288-293+313, China.

- [7] Li, S. K. (2017). Factors influencing the quality of mechanical grain harvesting in corn and development directions of grain harvesting technology in China (我国玉米机械粒收质量影响因素及粒收技术的发展方向). *Journal of Shihezi University (Natural Science Edition)*, Vol. 35(3), pp. 265-272, China.
- [8] Li, X., Du, Y., Yao, L., Wu, J., & Liu, L. (2021). Design and Experiment of a Broken Corn Kernel Detection Device Based on the Yolov4-Tiny Algorithm. *Agriculture*, Vol. 11, pp. 1238, China
- [9] Liu, Y. F., Zhong, Y. F., & Qin, Q. Q. (2018). Scene classification based on multiscale convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 56, pp. 7109–7121. China.
- [10] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp. 4510–4520, US.
- [11] Velesaca, H. O., Mira, R., Suárez, P. L., Larrea, C. X., & Sappa, A. D. (2020). Deep learning based corn kernel classification. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 294–302. USA
- [12] Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp. 7464–7475, China.
- [13] Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19, South Korea.
- [14] Wu, Y. S., Li, A. H., Wan, Y. M., & Meng, T. C. (2024). Research progress on online detection technology of tobacco based on machine vision (基于机器视觉的烟草在线检测技术研究进展). *Laser & Optoelectronics Progress*, Vol. 61(8), pp. 49-62, China.
- [15] Xu, T. C., Wu, M., He, D. X., Zheng, Z. A., Xu, H. H., & Bao, J. Q. (2021). Application of machine vision in agricultural engineering (机器视觉在农业工程中的应用). *Agricultural Engineering*, Vol. 11 (8) , pp. 40-48, China.
- [16] Xu, Y., Liu, L., Li, Z. Y., Gao, Z., & Li, X. Z. (2020). Corn variety recognition based on convolutional neural network (基于卷积神经网络的玉米品种识别). *Journal of Jiangsu Agricultural Sciences*, Vol. 36(01), pp. 18-23, China.
- [17] Yang, J. Y., Song, B., Luo, Y. J., Zhang, J., Liu, J., Yang, C., & Duan, J. M. (2018). Quality evaluation and preliminary screening of identification indicators for mechanical grain harvesting of different corn varieties (不同玉米品种机械粒收质量评价及其鉴定指标初步筛选). *Journal of Henan Agricultural Sciences*, Vol. 47(11), pp. 25-31, China.
- [18] Zhao, B., Li, X. L., Zhou, M. L., Song, B., Lei, E., Li, Z., Wu, Y. W., Yuan, J. C., & Kong, F. L. (2020). Current status and influencing factors of kernel breakage rate in mechanical grain harvesting of corn in Southwest China (西南玉米机械粒收籽粒破碎率现状及影响因素分析). *Acta Agronomica Sinica*, Vol. 46(1), pp. 74-83, China.
- [19] Zhao, C. J., Ma, C., Li, J., & Wang, X. M. (2025). Research status and prospects of rice mechanization technology in hilly and mountainous areas (丘陵山地水稻机械化技术研究现状与展望). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 41(1), pp. 1-11, China.
- [20] Zhu, X. L., Chi, R. J., Du, Y. F., Deng, X. J., Zhang, Z., & Dong, N. X. (2021). Design and experiment of intelligent control system for low-loss threshing of high moisture corn (高含水率玉米低损脱粒智能控制系统设计与试验). *Transactions of the Chinese Society for Agricultural Machinery*. Vol. 52(S1), pp. 9-18, China.