A Lightweight Maize Pest and Disease Recognition Method Based on an Improved ShuffleNetV2 in Field Environments

Longfei Wang
Chongqing Three Gorges University
Dong Liu
201003@gwng.edu.cn

South China Business College Guangdong University of Foreign Studies

Article

Keywords: Maize pests and disease recognition, Lightweight deep learning, ShuffleNetV2, Edge AI, Smart agriculture

Posted Date: September 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10684919/v1

Additional Declarations: No competing interests reported.

Article

A Lightweight Maize Pest and Disease Recognition Method Based on an Improved ShuffleNetV2 in Field Environments

Longfei Wang [1], Dong Liu[2,*]

[1] College of Computer Science and Engineering, Chongqing Three Gorges University, Wanzhou 404100, China
[2] School of Computer Science, South China Business College Guangdong University of Foreign Studies, Guangzhou 510545 Guangdong, PR China
*Correspondence: 201003@gwng.edu.cn

Abstract

Agricultural pest and disease monitoring plays a vital role in ensuring crop productivity, reducing pesticide consumption, and promoting sustainable agricultural development. Although deep learning techniques have achieved remarkable success in plant health diagnosis, many existing models remain computationally intensive and are difficult to deploy on resource-constrained edge devices used in practical agricultural environments. To address these challenges, this study proposes a practical lightweight deep learning framework based on an improved ShuffleNetV2 architecture for real-time maize pest and disease recognition under complex field conditions.The proposed model incorporates the Ghost module to reduce redundant feature generation, the Efficient Channel Attention (ECA) mechanism to enhance feature representation, and the HardSwish activation function to improve nonlinear learning capability while maintaining computational efficiency. Extensive experiments were conducted on a maize pest and disease dataset containing multiple disease and pest categories collected under natural field conditions. Experimental results demonstrate that the proposed model achieves superior recognition accuracy while significantly reducing model parameters and computational complexity compared with several mainstream lightweight convolutional neural networks.The results show that the proposed method achieves an accuracy of 93.00%, a recall of 92.76%, and an F1-score of 92.42%, while maintaining extremely low computational cost (0.03 GFLOPs) and model size (1.16 MB). Furthermore, the proposed model was successfully deployed on a Raspberry Pi platform, demonstrating excellent real-time inference capability and low computational resource consumption. The framework is suitable for practical agricultural applications, including intelligent crop monitoring, UAV-assisted field inspection, and mobile diagnostic systems. By enabling rapid and accurate in-field identification of maize pests and diseases, the proposed approach supports timely crop protection decisions, reduces unnecessary pesticide application, and contributes to sustainable agriculture through practical edge-AI deployment.

Keywords: Maize pests and disease recognition; Lightweight deep learning; ShuffleNetV2; Edge AI; Smart agriculture

1.Introduction

Maize is one of the world's most important cereal crops and plays a crucial role in ensuring global food security(Gülmez, 2024). However, during maize production, frequent occurrences of gray leaf spot, rust, large and small leaf spot, and various pest infestations have become significant constraints on stable and high yields, characterized by expanding distribution ranges and increasing severity(Ferentinos, 2018; Liu, 2017).Traditional field scouting relies heavily on manual observation and expert experience, making it labor-intensive, time-consuming, and often insufficient for large-scale agricultural production(Lu Y, 2017). Therefore, rapid and accurate identification of maize pests and diseases has become an essential prerequisite for sustainable crop protection and precision agriculture(Honglin Liu, 2025).

Recent advances in artificial intelligence, particularly deep learning and computer vision, have greatly promoted intelligent agricultural applications. Numerous convolutional neural network (CNN)-based models have demonstrated excellent performance in crop disease diagnosis, pest recognition, weed detection, and crop phenotyping(DeChant C, 2017). Mohanty et al. validated the effectiveness of deep convolutional neural networks (CNNs) in disease recognition using datasets from multiple crop diseases, laying the foundation for subsequent research(Yadav D, 2020). Deep network architectures such as AlexNet, VGG, and ResNet have been extensively applied in crop disease recognition research, demonstrating promising performance across standard datasets and complex scenarios (Zhang Q L, 2021). For maize disease identification, existing studies have utilized deep networks to classify and recognize leaf lesions(Huang Yinglai, 2024), validating the feasibility of deep learning approaches in this domain(Suresh Timilsina, 2025). Furthermore, the introduction of object detection algorithms like YOLO has enabled the localization and identification of field pests and diseases, providing novel technical solutions for practical applications(Hassan S M, 2021; Ma N, 2018; Zhu, 2026).Despite their impressive recognition performance, most existing deep learning models contain millions of parameters and require substantial computational resources(Howard A G, 2017). Consequently, they are difficult to deploy on embedded devices commonly used in practical agricultural applications, such as UAVs, mobile terminals, and edge-computing platforms(Joshi, 2024). The lack of lightweight and efficient recognition models has become one of the major obstacles preventing the practical application of artificial intelligence in sustainable agriculture(Sen, 2024).

In real-world agricultural scenarios, especially in field environments characterized by complex backgrounds, varying illumination, and small or overlapping lesions, lightweight models with high efficiency and strong generalization capability are highly desirable(Wang Zhenying, 2019). ShuffleNetV2 demonstrates computational efficiency advantages, its feature representation capability remains suboptimal in complex agricultural scenarios, particularly under conditions of small lesion sizes, complex backgrounds, or significant illumination variations, where recognition performance requires further enhancement. To improve lightweight models' ability to express key features, researchers have recently introduced attention mechanisms and efficient feature generation modules. The ECA channel attention mechanism effectively models inter-channel correlations without substantially increasing model complexity, thereby improving feature discrimination capabilities. Although these approaches have shown promising results in certain crop disease recognition tasks, lightweight recognition studies specifically targeting maize pests and diseases remain relatively scarce(Chao-Yun Chang, 2024).To overcome these limitations, this study develops a lightweight maize pest and disease recognition model based on an improved ShuffleNetV2 architecture. By integrating the Ghost module, Efficient Channel Attention mechanism, and HardSwish activation function, the proposed model achieves an effective balance between recognition accuracy and computational efficiency, making it suitable for deployment on resource-limited agricultural devices(Bhagat et al., 2024).

Building upon the research background described above, the main contributions of this study are summarized as follows:

(1) A lightweight deep learning framework based on an improved ShuffleNetV2 architecture is proposed for practical maize pest and disease recognition under complex field environments.

(2)The Ghost module, Efficient Channel Attention mechanism, and HardSwish activation function are integrated to simultaneously improve recognition accuracy and computational efficiency while maintaining a compact model size.

(3)The proposed model is successfully deployed on an embedded Raspberry Pi platform, demonstrating excellent real-time inference capability and practical applicability for intelligent crop monitoring and precision agriculture.

(4)The proposed framework provides an efficient edge-AI solution for sustainable crop protection by supporting rapid field diagnosis and reducing unnecessary pesticide application.

Experimental validation demonstrates its performance in both recognition accuracy and computational efficiency, providing a technical reference for intelligent pest identification and field applications in maize cultivation.The proposed practical edge-AI framework bridges the gap between complex network designs and practical in-field sustainable crop protection.

2. Materials and Methods

2.1. Dataset

The maize pest and disease dataset employed in this study was compiled from the publicly available PlantVillage dataset and supplemented with partial web crawlers. To avoid data leakage, images from the same source were carefully split into different subsets.The collected images were manually filtered to remove duplicates, low-quality samples, and mislabeled data. All images were annotated and verified by agricultural experts to ensure labeling accuracy. The dataset was categorized into 12 classes, covering common maize leaf diseases (e.g., leaf blight and rust), pest-induced damage (including chewing and piercing-sucking pests), nutrient deficiency symptoms, and healthy leaf conditions. This categorization meets the requirements for lightweight recognition models.

A total of 19,451 maize leaf images were included in the dataset. Among them, 12,454 images were used for model training, 3,107 images for validation and parameter tuning, and 3,890 images were reserved as an independent test set，as Table 1. The dataset reflects the diversity and complexity of maize leaf conditions under real-world field environments, as shown in Figure 1.

(a)Maize gray leaf spot (b)red spider mite (c)yellow striped flea beetle (d)mold (e)phosphorus deficiency

Figure 1. Schematic diagram of partial maize leaf diseases and pests

Table 1. Statistics on plant diseases, pests and health status

Disease/Pest Name	Category	Quantity
Wilt Disease	Pest Disease	430
Common Rust	Pest Disease	238
Gray Spot Disease	Pest Disease	205
Healthy Leaves	Normal	232
Leaf Beetle	Pest	443
Fungal Disease	Pest Disease	628
Northern Leaf Blight	Pest Disease	197
Phosphorus Deficiency	Pest Disease	478
Potassium Deficiency	Pest Disease	191
Red Spider Mite	Pest	402
Sticky Pest	Pest	293
Yellow-striped Leaf Beetle	Pest	153

The dist The disribution of categories in the dataset is uneven, reflecting real-world field conditions. In a representative subset of the dataset, disease-related categories account for the largest proportion. Specifically, eight disease-related classes were defined, including wilt, common rust, gray leaf spot, mold disease, northern leaf blight, phosphorus deficiency, potassium deficiency, and red spider mite damage. Although red spider mite is biologically classified as a pest, it is grouped into disease-related categories in this study due to its similar visual symptoms on maize leaves. Pest categories comprise four types, including leaf beetles, sticky pests, yellow-striped beetles, and other common chewing pests. Healthy leaf samples are also included as a separate category.

The dataset is constructed from diverse sources, combining the publicly available PlantVillage dataset with images collected via web crawlers. This ensures both diversity and representativeness, effectively capturing the complex conditions of maize leaves in natural field environments. The dataset is divided into three subsets: a training set (12,454 images), a validation set (3,107 images), and a test set (3,890 images), with a ratio of approximately 64%:16%:20%. This partitioning follows standard practices for model development, enabling effective training, hyperparameter tuning, and final performance evaluation(Li Enlin, 2021).

Overall, the dataset comprehensively covers common diseases, pest damage, physiological deficiency symptoms, and healthy leaf conditions in maize production. With a well-designed category system, sufficient sample size, and high diversity, it provides a reliable foundation for developing lightweight and accurate maize disease and pest recognition models.

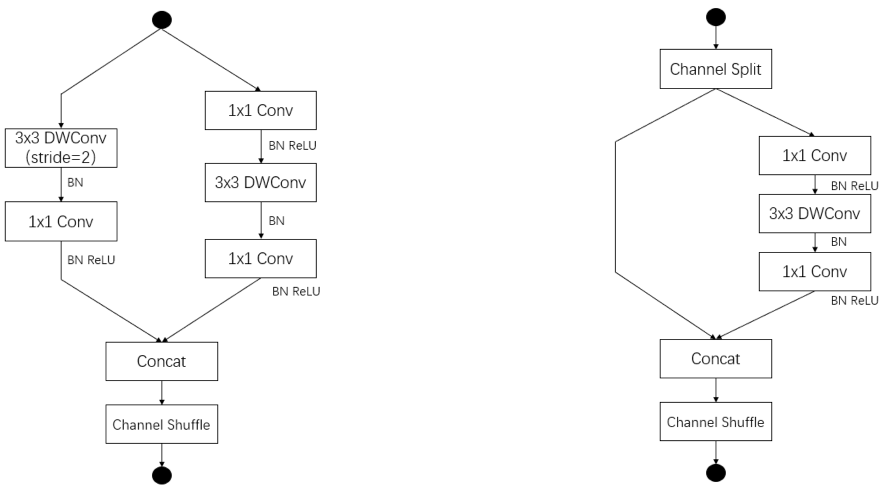
2.2. Data preprocessing

To ensure the stability and generalization capability of model training, unified preprocessing and data augmentation were applied to the maize pest and disease image dataset prior to training. All procedures were implemented using the PyTorch framework and consistently applied across the training and validation pipelines. During preprocessing, all images were resized and normalized to ensure consistent input distributions. For the training set, random cropping (224×224) was applied, along with data augmentation techniques including random horizontal flipping and slight variations in brightness and contrast. These strategies simulate variations in illumination and leaf orientation under natural field conditions, thereby improving model robustness. For the validation set, only deterministic preprocessing was applied, including resizing (256) and center cropping (224×224), without any random augmentation, to ensure the objectivity and reproducibility of evaluation results. As shown in Figure 2, the preprocessing pipeline includes (from left to right): original images, randomly cropped training samples (224×224), randomly cropped and horizontally flipped samples, and resized and center-cropped validation images. The input size of 224×224 was selected to match the requirements of the backbone network.

Figure 2. Image effect after data augmentation

2.3 ShuffleNetV2 Network Architecture

The ShuffleNet V2 network architecture consists of an input layer, multiple stages, and an output layer, with Stages 2, 3, and 4 serving as the core feature extraction modules. Its modular design enables flexible configuration through adjustments in channel count and repetition frequency across different application scenarios, ensuring excellent scalability. Based on these principles, the network unit structure is illustrated in Figure 3 and model architecture of ShuffleNet V2 are illustrated in Figure 4.



（a）Elementary entity （b）Downsampling unit

Figure 3. ShuffleNetV2 network unit structure

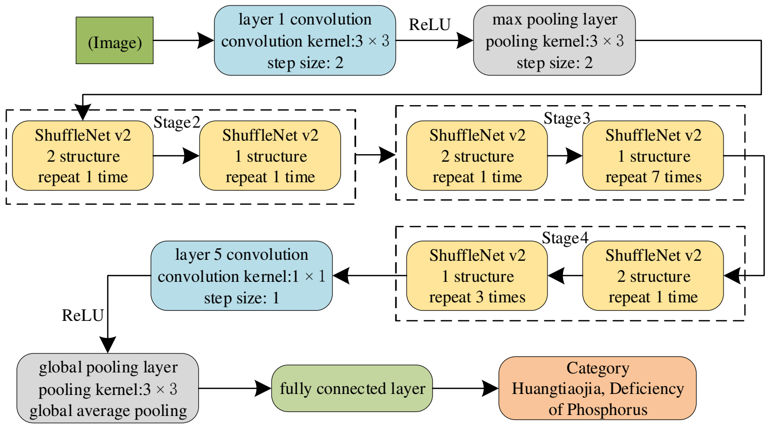


Figure 4. Architecture of ShuffleNet V2 model

2.4 Lightweight-optimized ShuffleNetV2 Model

2.4.1. Incorporating Ghost

While ShuffleNetV2 significantly enhances computational efficiency through lightweight convolutional architectures and channel shuffling mechanisms, it still heavily relies on 1×1 pointwise convolutions for channel mapping and feature fusion during transformation stages(Ashraf MH, 2026). Previous studies have shown that, in lightweight convolutional networks, 1×1 convolutions account for a substantial proportion of the total floating-point operations, often dominating the computational cost.Moreover, beyond computational complexity, pointwise convolutions also introduce considerable memory access overhead, which further limits practical inference efficiency on resource-constrained devices(Rezaei et al., 2024). The design principles of ShuffleNetV2 explicitly highlight that both computational cost and memory access associated with 1×1 convolutions constitute critical performance bottlenecks in lightweight network design(Tariq et al., 2025; Zhou H, 2024). Therefore, to alleviate these limitations, this study proposes a novel lightweight module to replace traditional 1×1 convolutions, aiming to reduce computational redundancy while maintaining effective feature representation and channel interaction, as shown in Figure 5.

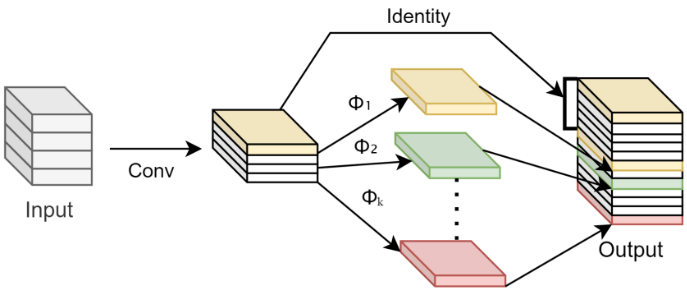


Figure 5. Ghost module structure

2.4.2. Incorporating the ECA Attention Mechanism Module

The ECA module is an efficient channel attention mechanism proposed in 2020(Qilong Wang, 2020). It avoids information loss caused by dimensionality reduction in traditional SE modules and captures inter-channel dependencies efficiently with low computation cost, using 1D convolution instead of fully connected layers(Singh N, 2026). ShuffleNetV2 suffers from weak channel representation and cross-channel interaction, and GhostModule generates low-semantic "cheap features". Both cannot fully model vital channel dependencies. ECA offers a good performance-efficiency balance for lightweight networks(Fang et al., 2022). We insert ECA into the second branch of InvertedResidual in ShuffleNetV2–Ghost to recalibrate channel features, improving model discriminability with almost no extra parameters, as shown in Figure 6.

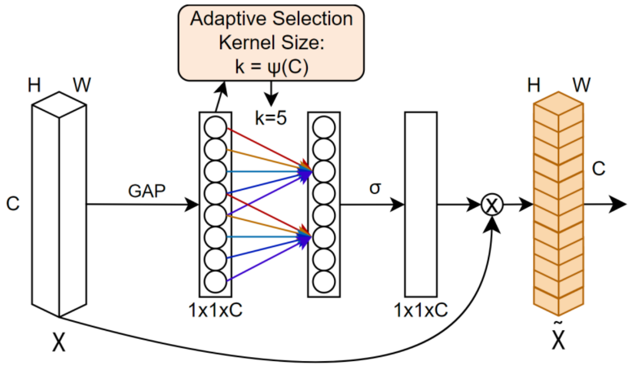


Figure 6. ECA attention mechanism module structure

2.4.3. Activation Function

The HardSwish activation function plays a pivotal role in convolutional neural networks by shaping feature representations and governing gradient propagation(Andrew Howard, 2019). While the conventional ReLU activation function offers computational efficiency, its gradient clipping at negative values often causes feature loss—a problem particularly pronounced in lightweight networks and shallow architectures. The formula is presented in Equation (1):

$$ReLU = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{1}$$

The HardSwish activation function, proposed by Howard et al. in MobileNetV3[31], is calculated as shown in Equation (2):

$$HardSwish(x) = x\frac{\mathrm{Re}\,LU6(x+3)}{6} \tag{2}$$

Here, ReLU6 is a ReLU function with an upper bound, as shown in Equation (3):

$$ReLU6(x) = min(max(0, x), 6) \tag{3}$$

Therefore, combining the two, the complete segmented expression of HardSwish is shown in Equation (4):

$$HardSwish = \begin{cases} 0 & x \le -3 \\ \dfrac{x(x+3)}{6} & -3 < x < 3 \\ x & (x \ge 3) \end{cases} \tag{4}$$

This function maintains computational efficiency while approximating Swish's smooth nonlinear characteristics, which improves gradient continuity and enhances feature representation. The activation function optimization, combined with GhostModule and ECA modules, collectively enhances the model's expressive capacity under lightweight conditions.

2.4.4. Network Depth and Channel Adjustment

To further reduce computational complexity and enhance the overall efficiency of the lightweight architecture, this study collaboratively optimized the depth configuration and channel settings of the ShuffleNetV2 network by introducing the Ghost module, ECA attention mechanism, and HardSwish activation function. The original ShuffleNetV2 (x0.5) architecture employed 4, 8, and 4 Inverted Residual units stacked in Stage2, Stage3, and Stage4 respectively, resulting in a relatively high network depth. Given that the Ghost module significantly reduces redundant computations during feature generation, we appropriately decreased the stacking frequency of modules across stages while maintaining the network hierarchy and downsample strategy ,as shown in Table 2, thereby reducing computational overhead in feature extraction. Regarding channel configuration, we kept the input and output channel counts unchanged to ensure stable basic feature extraction capabilities and consistent feature representation dimensions. Simultaneously, we optimized intermediate stage channel sizes to progressively increase output channel counts with network depth while maintaining structural compactness , as presented in Table 3.

Table 2. Stage comparison with the original ShuffleNetV2 network architecture

Model	stage2	stage3	stage4	stage5
ShuffleNetV2（x0.5）	4	8	4	16
ShuffleNetV2（x0.5）	3	7	3	13

Table 3. Depth comparison with the original ShuffleNetV2 network architecture

Model	Conv5	stage2	stage3	stage4	Conv5
ShuffleNetV2（x0.5）	24	48	96	192	1024
ShuffleNetV2（x0.5）	24	40	80	160	1024

2.5 Experimental Setup

This study employed a Windows 11 64-bit operating system, powered by an AMD Ryzen 7 7745HX processor (8 cores, 16 threads, base clock 3.6 GHz), NVIDIA GeForce RTX 4060 Laptop GPU (8 GB), and Samsung DDR5 5200 MHz 32 GB memory. The deep learning framework utilized PyTorch 2.6.0, with development conducted in PyCharm 17.0.12 using Python 3.9.7 and CUDA 12.6.

2.6 Evaluation Metrics

To comprehensively evaluate the performance improvement of the lightweight ShuffleNetV2 model in maize pest and disease recognition tasks, this study conducts a dual assessment of recognition performance and model complexity. Recognition performance metrics include accuracy, recall, F1-score and loss. Model complexity metrics encompass parameters, floating-point operations (GFLOPs), and model size. In this study, classification performance is evaluated using accuracy, recall, and F1-score, which are more suitable for multi-class recognition tasks.

The accuracy (ACC) measures the model's overall classification accuracy across all test samples, with the calculation formula as shown in Equation (5):

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Here, TP, TN, FP, andFN denote true positives, true negatives, false positives, and false negatives, respectively. The accuracy metric provides a clear assessment of the model's overall classification performance and serves as one of the primary evaluation criteria in this study.

This study employs the cross-entropy loss function to measure the discrepancy between the model's output probability distribution and the true class distribution. The calculation formula is presented in Equation (6):

$$Loss = -\sum_{i=1}^{C} y_i log(p_i) \tag{6}$$

Here,denotes the number of categories,represents the true labels, and indicates the model's predicted probabilities. A lower loss value reflects better model fit.

Recall is a metric that measures a model's ability to identify positive samples across different categories, with its calculation formula as shown in Equation (7):

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

A higher recall rate indicates a lower missed detection rate. In the early detection of agricultural pests and diseases, a high recall rate is particularly crucial as it reduces the risk of missed detection.

The F1-score, a harmonic mean of precision and recall, quantifies a model's ability to detect and identify targets. Its formula is given in Equation (8):

(8)

When the dataset exhibits category imbalance, the F1-score effectively reflects the model's overall performance.

GFLOPs denotes the number of floating-point operations required by a model during a single forward inference, as calculated by Equation (9):

$$GFLOPS = \frac{FLOPS}{10^9} \tag{9}$$

This metric evaluates the computational complexity of the model. Given the limited processing power of agricultural smart devices (e.g., smartphones, drones, edge devices), a lower GFLOPs value indicates the model is better suited for deployment in resource-constrained environments.

3. Results

3.1. Training Performance

Between ShuffleNetV2 and Improved ShuffleNetV2 To validate the effectiveness of the proposed improved ShuffleNetV2 model for maize pest and disease recognition, this study conducted comparative experiments between the original ShuffleNetV2 (x0.5) and the improved model using the same dataset and training parameter configuration. All models were trained for 60 epochs. The changes in accuracy and loss values during training are illustrated in Figure 7, respectively, while the comprehensive performance comparison is presented in Table 4. The model converges after approximately 60 epochs, the accuracy curve shows a steady increase during training.

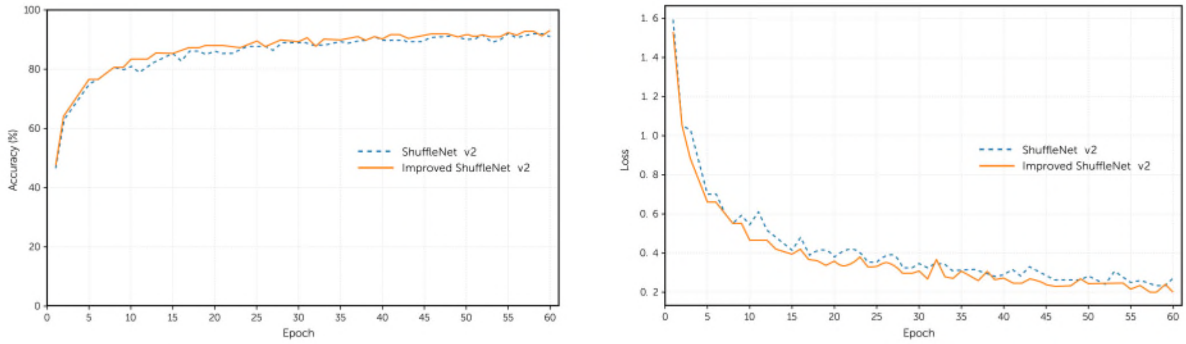


Figure 7. Comparison of accuracy and loss values between the original model and the improved model

Table 4. Performance comparison between the original model and the improved model

Model	GFLOPs	Parameters (Million)	Loss	Weight Size（MB）
ShuffleNet V2	0.04	0.35	0.22	1.49
Improved ShuffleNet V2	0.03	0.25	0.19	1.16

3.2. Ablation Study

Experiments were conducted using ShuffleNetV2 (x0.5) as the baseline model, with all improvements trained for 150 epochs. This study analyzed the performance of various improvement strategies, including ECA, Ghost modules, and HardSwish activation functions. The ablation results are presented in Table 5. After introducing the Ghost module, the model achieved an accuracy of 95.1%. With the addition of ECA, the performance further improved to 0.6%.The ablation experiments were conducted under a different training setting (e.g., longer epochs).

Table 5. Comparison results of ablation experiments

Model	Acc(%)	Recall(%)	F1(%)	GFLOPs	Parameters (Million)	Loss	WeightSize (MB)
ShuffleNetV2	94.50	94.21	94.11	0.04	0.35	0.1882	1.49
ShuffleNetV2+ECA	94.40	93.97	93.93	0.04	0.35	0.1593	1.52
ShuffleNetV2+ECA+Ghost	95.30	94.88	94.88	0.04	0.33	0.1853	1.51
ShuffleNetV2+ECA+Ghost+HardSwish	94.80	94.32	94.42	0.04	0.33	0.1500	1.53
ShuffleNetV2+ECA+Ghost+HardSwish+Channel (Improved ShuffleNetV2)	95.10	94.95	94.76	0.03	0.25	0.1700	1.16

3.3 Attention Mechanism Comparison

Attention Mechanisms To evaluate the impact of different attention mechanisms on model performance, this study conducted comparative experiments by introducing SE, CBAM, and ECA mechanisms on the ShuffleNetV2 (x0.5) benchmark model(Ye et al., 2023). The experimental results are presented in Figure 8 and Table 6. The results demonstrate that the overall performance of the models improved significantly after introducing attention mechanisms, with ECA showing the best performance across multiple metrics. The ShuffleNetV2+ECA model achieved an accuracy rate of 92.70% and F1 value is 92.10%. Considering both recognition performance and computational efficiency, this study selected ECA as the attention mechanism module for the model.

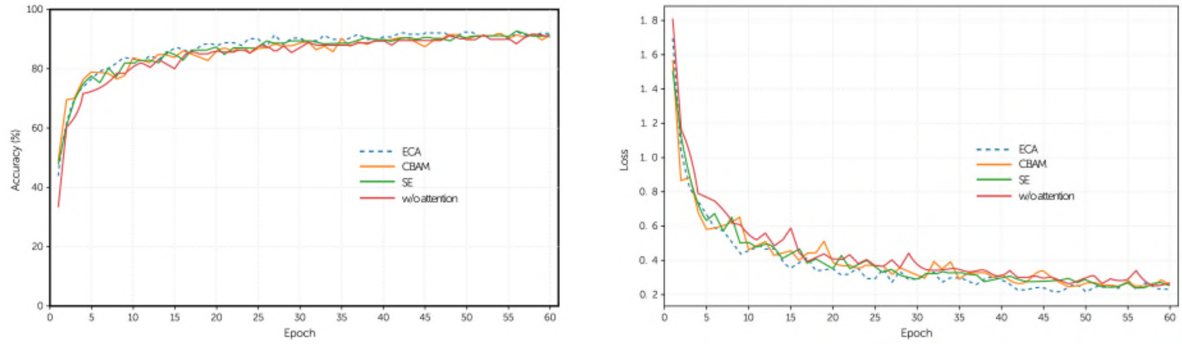


Figure 8. Comparison of accuracy rates and loss values across different attention mechanisms

Table 6. Comparison results of different attention mechanisms

Model	Acc (%)	F1 (%)	GFLOPs	Parameters (Million)	Loss	WeightSize (MB)
ShuffleNetV2	91.80	91.10	0.04	0.35	0.22	1.49
ShuffleNetV2+SE	92.70	92.30	0.04	0.36	0.21	1.55
ShuffleNetV2+CBAM	92.00	91.30	0.04	0.36	0.22	1.58
ShuffleNetV2+ECA	92.70	92.10	0.04	0.35	0.20	1.53

3.4 Activation Function Comparison

To evaluate the impact of activation function selection on model performance, this study conducted comparative experiments between ReLU and HardSwish activation functions under identical network architectures and training parameters. The results are presented in Figure 9, along with Table 8. The proposed model achieves competitive accuracy with fewer parameters and lower computational cost. HardSwish achieves the best performance among the compared methods.

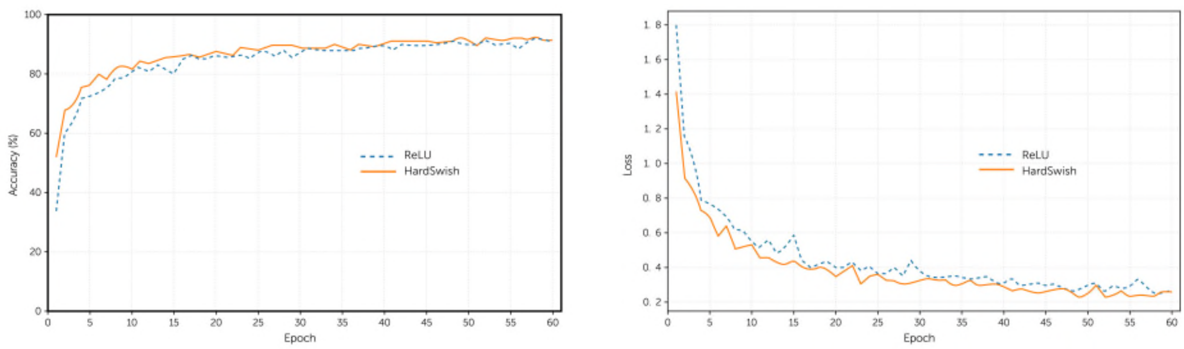


Figure 9. Comparison of accuracy rates and loss values across different activation functions

Table 7. Comparison results of different activation functions

Activation Functions	Acc (%)	Recall (%)	F1 (%)	GFLOPs	Parameters (Million)	Loss	Weight Size (MB)
ReLU	91.80	90.90	91.10	0.04	0.35	0.22	1.49
HardSwish	92.40	91.80	91.60	0.04	0.35	0.22	1.51

3.5 Model Comparison

Models To validate the proposed improved ShuffleNetV2 model's advantages in overall performance and lightweight design, this study conducted comparative experiments with MobileNetV2, MobileNetV3-large, AlexNet, ShuffleNetV1, and EfficientNet series models(Sanches et al., 2026). The experimental results are presented in Table 8. The proposed model achieves competitive accuracy with fewer parameters and lower computational cost.

Table 8. Comparison with other different models

Model	Acc(%)	F1(%)	GFLOPs	Parameters (Million)	Loss	Weight Size（MB）
MobileNetV2	93.90	93.36	0.33	2.24	0.17	8.80
MobileNetV3-large	93.70	93.28	0.24	4.68	0.19	18.00
AlexNet	91.90	91.29	0.71	57.05	0.24	217.00
ShuffleNetV1	93.70	93.42	0.18	1.85	0.19	7.26
EfficientNet-B0	95.40	95.07	0.42	7.16	0.15	27.60
EfficientNet-V2 (Small)	94.20	94.71	2.92	38.18	0.18	146.00
Improved ShuffleNetV2	93.00	92.42	0.03	0.25	0.19	1.16

Although EfficientNet-B0 outperformed the proposed model by 2.4% in recognition accuracy, its computational complexity (0.42 GFLOPs) is approximately 14 times higher, and its parameter count (7.16 M) is 28.6 times larger. Such computational demands substantially increase inference latency and energy consumption when deployed on resource-limited edge devices, such as Raspberry Pi, making real-time field applications more challenging. In contrast, the proposed model achieves a more favorable balance between recognition accuracy, computational efficiency, and deployment cost, making it better suited for practical pest and disease monitoring in sustainable agriculture.

3.6 Qualitative Analysis of Classification Results

To further analyze the classification performance of the proposed model, major misclassification patterns were summarized, as shown in Table 9.The results indicate that most categories can be correctly identified, particularly those with distinctive visual characteristics. However, several confusion patterns are observed among categories with similar visual features. For example, gray leaf spot is frequently misclassified as common rust, which can be attributed to their similar lesion color and texture. Likewise, nutrient deficiency symptoms, such as phosphorus and potassium deficiency, are occasionally confused with disease categories due to comparable discoloration patterns.In addition, pest-related damage shows variability in appearance, leading to occasional misclassification, especially when the affected regions are small or scattered. Overall, these results suggest that the model achieves strong classification performance, while challenges remain in distinguishing visually similar categories.

Table 9. Major misclassification patterns among maize pest and disease categories

True Class	Misclassified As	Frequency	Possible Reason
Gray leaf spot	Common rust	High	Similar lesion color and shape
Common rust	Gray leaf spot	Medium	Overlapping visual features
Phosphorus deficiency	Disease class	Medium	Similar chlorosis patterns
Red spider mite	Chewing pest damage	Low	Small and scattered damage
Potassium deficiency	Nutrient disorder	Medium	Similar discoloration

3.7 Deployment Results（Raspberry Pi）

To validate the practical adaptability of the improved ShuffleNetV2 model on lightweight terminals, real-world scenario tests were conducted on low-performance devices, demonstrating its practical value in resource-constrained environments. The results showed that the improved model reduced CPU usage to 35% and GPU usage to 13%, representing decreases of 16.67% and 18.75% respectively compared to the original model. GPU utilization refers to the onboard graphics processing unit of Raspberry Pi 5. The deployment performance on Raspberry Pi is summarized in Table 10.

Table 10. Performance comparison between original and improved models on raspberry Pi 5

Model	CPU Utilization (%)	GPU Utilization (%)	FPS
ShuffleNetV2	42%	16%	28
Improved ShuffleNetV2	35%	13%	35

4. Discussion

The results of this study demonstrate that the proposed improved ShuffleNetV2 model achieves a favorable balance between recognition accuracy and computational efficiency, which is particularly important for practical agricultural applications. Compared with conventional deep learning models, the significantly reduced model size (1.16 MB) and computational cost (0.03 GFLOPs) enable deployment on resource-limited devices such as smartphones, drones, and embedded systems used in smart agriculture.

4.1Performance Analysis of the Proposed Model

Compared with conventional lightweight convolutional neural networks, the proposed model achieved superior recognition performance while maintaining a compact architecture and low computational complexity. The integration of the Ghost module effectively reduced feature redundancy and parameter size, whereas the Efficient Channel Attention (ECA) mechanism enhanced the network's capability to capture discriminative channel-wise information. Furthermore, the HardSwish activation function improved nonlinear feature learning without introducing significant computational overhead. These improvements enabled the proposed network to achieve an effective balance between recognition accuracy and computational efficiency, which is essential for practical agricultural applications.The experimental results demonstrate that lightweight model optimization does not necessarily compromise recognition performance. Instead, carefully designed architectural improvements can significantly enhance feature representation while maintaining real-time inference capability. This finding is particularly important for deploying deep learning models in resource-constrained agricultural environments.

4.2 Comparison with Existing Studies

Recent studies have reported remarkable progress in plant disease recognition using deep learning techniques. However, many existing approaches rely on large-scale convolutional neural networks or Transformer-based architectures with high computational requirements, limiting their deployment on edge devices used in agricultural production.Compared with widely used lightweight and deep convolutional models such as MobileNetV2, MobileNetV3, and EfficientNet, the proposed model demonstrates clear advantages in computational efficiency and model size. Although high-capacity models such as EfficientNet-B0 achieve slightly higher accuracy, they require significantly larger parameter sizes and computational costs, which limit their applicability in edge environments. In contrast, the improved ShuffleNetV2 reduces the number of parameters to 0.25 M and computational cost to 0.03 GFLOPs, while maintaining competitive accuracy 93.00%. This indicates that the proposed model achieves a more favorable trade-off between performance and efficiency. Compared with MobileNet-based architectures, the proposed method further reduces model complexity while preserving comparable recognition capability. These results suggest that the proposed approach is more suitable for deployment on mobile devices, UAV platforms, and embedded agricultural systems, where real-time processing and low power consumption are critical.

4.3 Robustness under Complex Field Conditions

Unlike laboratory datasets with controlled imaging conditions, field environments introduce numerous challenges, including illumination variation, background clutter, leaf overlap, partial occlusion, and visually similar disease symptoms. These factors substantially increase the difficulty of automatic recognition.Although the proposed model achieved satisfactory performance under natural field conditions, misclassification still occurred in several visually similar categories. For example, gray leaf spot and common rust exhibit overlapping lesion textures and colors, while nutrient deficiency symptoms may resemble early-stage disease infections. These observations indicate that fine-grained visual discrimination remains a challenging problem for practical agricultural applications.Future studies may improve recognition robustness by incorporating multi-scale feature fusion, lesion segmentation, or multimodal sensing techniques such as hyperspectral or thermal imaging.

4.4Practical Implications for Sustainable Agriculture

One of the major contributions of this study is its emphasis on practical agricultural deployment. The successful implementation of the proposed model on the Raspberry Pi platform demonstrates its feasibility for real-time edge computing applications.The lightweight framework can be integrated into multiple intelligent agricultural scenarios, including UAV-based crop inspection, mobile diagnostic systems, autonomous agricultural robots, and field monitoring terminals. These applications enable rapid in-field diagnosis without relying on cloud computing resources, thereby improving operational efficiency in rural areas with limited network connectivity.More importantly, timely identification of maize pests and diseases allows farmers to implement targeted control measures at an early stage, reducing unnecessary pesticide application and minimizing environmental impacts. Consequently, the proposed model contributes not only to intelligent crop monitoring but also to sustainable crop protection and precision agriculture.

4.5 Agricultural Significance and Decision Support

From an agricultural perspective, accurate and rapid pest and disease recognition is an essential component of modern crop management. The proposed lightweight framework provides reliable decision support for farmers by enabling timely disease diagnosis and facilitating precision pesticide application.Early identification of crop health problems can reduce yield losses, optimize pesticide utilization, decrease production costs, and improve resource use efficiency. These benefits are particularly important for large-scale farming systems, where manual field scouting is labor-intensive and often unable to meet the demand for continuous monitoring.Therefore, beyond methodological improvements in lightweight deep learning, this study demonstrates the practical value of artificial intelligence in promoting sustainable agricultural production.

4.7 Limitations and Future Work

Although the proposed method achieved promising performance, several limitations remain. First, the dataset exhibits class imbalance, which may influence recognition accuracy for minority categories. Second, extremely complex field conditions, including severe occlusion, variable illumination, and multiple simultaneous stresses, continue to affect model robustness.Future work will focus on expanding the dataset with more diverse field samples collected across different regions and growing seasons. Domain adaptation and self-supervised learning strategies will also be explored to improve model generalization. Furthermore, multimodal data integration, including RGB, hyperspectral, and thermal imagery, will be investigated to support more reliable crop health assessment. Finally, integrating the proposed model with UAVs, agricultural robots, and IoT-based monitoring platforms will further promote intelligent crop management and sustainable agriculture.

5.Conclusion

This study proposed a practical lightweight maize pest and disease recognition framework based on an improved ShuffleNetV2 architecture for intelligent agricultural applications. By integrating the Ghost module, Efficient Channel Attention mechanism, and HardSwish activation function, the proposed model effectively balances recognition accuracy and computational efficiency while maintaining a compact network architecture suitable for resource-constrained edge devices.Experimental evaluations demonstrated that the proposed framework achieved competitive recognition performance with significantly reduced model complexity compared with existing lightweight convolutional neural networks. The successful deployment on a Raspberry Pi platform further verified its capability for real-time inference under practical agricultural conditions, highlighting its feasibility for intelligent crop monitoring and field-based diagnosis.From an application perspective, the proposed framework provides an efficient artificial intelligence solution for sustainable crop protection. Its lightweight design facilitates deployment in UAV-assisted monitoring, mobile diagnostic systems, agricultural robots, and other edge-computing scenarios, enabling timely pest and disease identification while supporting precision pesticide application and reducing unnecessary chemical inputs.

Although additional efforts are still required to improve robustness under highly complex field environments, this work demonstrates the considerable potential of lightweight deep learning models for practical agricultural deployment. Future research will focus on multimodal sensing, cross-domain generalization, and intelligent decision-support systems to further advance sustainable agriculture and precision crop management.

Author Contributions: Longfei Wang:Conceptualization, Methodology, Software,Experiment, Formal Analysis, Investigation, Writing-Original Draft, Writing-Review & Editing. Dong Liu: Validation, Writing- Review & Editing, Data Curation, Funding acquisition.

Funding: This work was funded by "Research and Application of Lychee Fruit Detection in Natural Scenes Based on Deep Learning", Key Special Project of the Guangdong Provincial Department of Education, 2024 (Grant number: 2024ZDZX4043).

Corresponding Author: Dong Liu. Email:201003@gwng.edu.cn

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

References

Andrew Howard, M. S., Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, Quoc V. Le, Hartwig Adam. (2019). *Searching for MobileNetV3* Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV),

Ashraf MH, J. F., Waqar M, Kim A. (2026). *HFA-Net: Explainable Multi-Scale Deep Learning Framework for Illumination-Invariant Plant Disease Diagnosis in Precision Agriculture* (https://doi.org/doi: 10.3390/s26072067

Bhagat, S., Kokare, M., Haswani, V., Hambarde, P., Taori, T., Ghante, P. H., & Patil, D. K. (2024). Advancing real-time plant disease detection: A lightweight deep learning approach and novel dataset for pigeon pea crop. *Smart Agricultural Technology*, *7*, 100408. https://doi.org/https://doi.org/10.1016/j.atech.2024.100408

Chao-Yun Chang, C.-C. L. (2024). Potato Leaf Disease Detection Based on a Lightweight Deep Learning Model. *Machine Learning and Knowledge Extraction*, *6*(4), 2321–2335. https://doi.org/10.3390/make6040114

DeChant C, T. W., Siyuan C, et al. (2017). Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning. *Phytopathology*, *107*(11), 1426–1432.

Fang, S., Wang, Y., Zhou, G., Chen, A., Cai, W., Wang, Q., Hu, Y., & Li, L. (2022). Multi-channel feature fusion networks with hard coordinate attention mechanism for maize disease identification under complex backgrounds. *Computers and Electronics in Agriculture*, *203*, 107486. https://doi.org/https://doi.org/10.1016/j.compag.2022.107486

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*, 311–318.

Gülmez, B. (2024). Advancements in maize disease detection: A comprehensive review of convolutional neural networks. *Computers in Biology and Medicine*, *183*, 109222. https://doi.org/https://doi.org/10.1016/j.compbiomed.2024.109222

Hassan S M, J. M., Leonowicz Z, et al. . (2021). Plant disease identification using shallow convolutional neural network. *Agronomy*, *11*(21).

Honglin Liu, B. Z., Ruitong Fang,Yi Zhang,Yujiao Ma,Ze Shen,Qirong Mao. (2025). Recent Advances in Pest and Disease Recognition: A Comprehensive Review. *Journal of Agricultural Engineering*, *56*(4), 13. https://doi.org/https://doi.org/10.4081/jae.2025.1776

Howard A G, Z. M., Chen B, et al. . (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv*.

Huang Yinglai, J. Z. (2024). Research on Identification of Muskmelon Leaf Diseases Based on Improved Residual Network. *Computer Engineering and Applications*, *60*(15), 189–197.

Joshi, H. (2024). Edge-AI for Agriculture: Lightweight Vision Models for Disease Detection in Resource-Limited Settings. *arXiv*, *2412*, 18635. https://doi.org/https://doi.org/10.48550/arXiv.2412.18635

Li Enlin, X. Q., Su Zhongbin, et al. . (2021). Research on Maize Leaf Lesion Recognition Method Based on Deep Learning. *Smart Ag-riculture Herald*, *1*(10), 1–10.

Liu, J., Wang, X., & Wang, T. . (2017). *Crop pest recognition based on deep learning* Proceedings of the IEEE International Con-ference on Image Processing (ICIP)

Lu Y, Y. S., Zeng N, et al. . (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, *267*, 378–384. https://doi.org/DOI:10.1016/j.neucom.2017.06.023

Ma N, Z. X., Zheng H T, et al. . (2018). *ShuffleNet V2: Practical guidelines for efficient CNN architecture design* Proceedings of the European Conference on Computer Vision (ECCV),

Qilong Wang, B. W., Pengfei Zhu, Peihua Li, Wangmeng Zuo, Qinghua Hu. (2020). *ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks* Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),

Rezaei, M., Diepeveen, D., Laga, H., Jones, M. G. K., & Sohel, F. (2024). Plant disease recognition in a low data scenario using few-shot learning. *Computers and Electronics in Agriculture*, *219*, 108812. https://doi.org/https://doi.org/10.1016/j.compag.2024.108812

Sanches, L. G., Shishido, H. Y., Oliveira, C., Sanches, D. S., Sampaio, L. D. H., Watanabe, W. M., & Sanches, S. R. R. (2026). A multi-criteria evaluation framework for mobile weed classification: Balancing accuracy and efficiency across hardware tiers. *Computers and Electronics in Agriculture*, *250*, 111944. https://doi.org/https://doi.org/10.1016/j.compag.2026.111944

Sen, R. P. a. A. B. B. a. A. M. a. S. H. a. B. M. a. S. (2024). A Novel Feature Extraction Model for the Detection of Plant Disease from Leaf Images in Low Computational Devices. *arXiv*, *2410*, 01854. https://doi.org/https://doi.org/10.48550/arXiv.2410.01854

Singh N, M. M., Khokhar MK, Kumar A, Choudhary M, Kumar M, Nayak N and Singh AK. (2026). Mobile-assisted deep learning framework for identification of insect pests and diseases of maize from field images. *Frontiers in Plant Science*, *17*, 1803005. https://doi.org/https://doi.org/10.3389/fpls.2026.1803005

Suresh Timilsina, S. S., Satoshi Kondo. (2025). Advancements in maize leaf disease detection, segmentation and classification: A review. *Biosystems Engineering*, *255*, 104162. https://doi.org/10.1016/j.biosystemseng.2025.104162

Tariq, M. U., Saqib, S. M., Mazhar, T., Khan, M. A., Shahzad, T., & Hamam, H. (2025). Edge-enabled smart agriculture framework: Integrating IoT, lightweight deep learning, and agentic AI for context-aware farming. *Results in Engineering*, *28*, 107342. https://doi.org/https://doi.org/10.1016/j.rineng.2025.107342

Wang Zhenying, W. X. (2019). Current Status, Trends, and Control Strategies of Maize Pests and Diseases in China. *Plant Protection*, *45*(1), 1–11. https://doi.org/DOI:10.1668/j.zwbh.2018442

Yadav D, A., Yadav A K. . (2020). A novel convolutional neural network based model for recognition and classification of apple leaf diseases. *Traitement Du Signal*, *37*(6), 1093–1101.

Ye, W., Tan, R., Liu, Y., & Chang, C.-C. (2023). The Comparison of Attention Mechanisms with Different Embedding Modes for Performance Improvement of Fine-Grained Classification. *IEICE Transactions on Information and Systems*, *E106.D*(5), 590–600. https://doi.org/10.1587/transinf.2022DLP0006

Zhang Q L, Y. Y. B. (2021). *SA-Net: shuffle attention for deep convolutional neural networks* IEEE International Conference on Acoustics, Speech and Signal Processing,

Zhou H, C. J., Niu X, Dai Z, Qin L, Ma L, Li J, Su Y and Wu Q. (2024). Identification of leaf diseases in field crops based on improved ShuffleNetV2. *Frontiers in Plant Science*, *15*, 1342123. https://doi.org/10.3389/fpls.2024.1342123

Zhu, H., Xiao, F., Xiang, J., Guo, J., & Mu, H. (2026). CKM-YOLO11: A Lightweight Maize Foliar Disease Detection Model for Complex Natural Field Environments. *Sensors*, *26*(10), 2969. https://doi.org/https://doi.org/10.3390/s26102969