

Classification of wild mushrooms based on improved ShuffleNetV2

Xingmei Xu¹, Dawei Yang¹, Qiqi Wei², Jinying Li^{3*}, Jian Zhang^{4,5*}

- (1. College of Information Technology, Jilin Agricultural University, Changchun 130118, China;
2. College of Computer Engineering, Qingdao City University, Qingdao 266106, China;
3. College of Horticulture, Jilin Agricultural University, Changchun 130118, China;
4. Faculty of Agronomy, Jilin Agricultural University, Changchun 130118, China;
5. Department of Biology, University of British Columbia, Okanagan, Kelowna, BC V1V 1V7, Canada)

Abstract: This study introduced an improved CHE_ShuffleNetV2 model based on ShuffleNetV2 to address the classification challenge of wild mushrooms in a complex environment. The model incorporated a Cross Stage Partial (CSP) structure to simplify its complexity. Furthermore, it adopted Hybrid Dilated Convolution (HDC) to replace conventional convolution, enhancing the model's recognition accuracy by expanding its receptive field. In addition, the ECA module was integrated to enhance the focus of the model on crucial feature information. The Hardswish activation function was employed instead of the ReLU activation function to reduce the number of parameters. The experimental results demonstrated that the enhanced model achieved improved accuracy, precision, recall, and F1-Score of 95.02%, 95.19%, 94.56%, and 94.00%, respectively, representing improvements of 2.81%, 3.82%, 3.08%, and 3.65%, correspondingly, over the original model. The enhanced model also reduced the parameters and FLOPs to 0.933 M and 104.08 M, respectively, representing reductions of 26.13% and 30.42% over the original model. Compared with commonly used lightweight models such as EfficientNet, DenseNet, and MobileNetV2, the CHE_ShuffleNetV2 model showed superior performance in solving the wild mushroom classification problem in complex environments, exhibiting its suitability for deployment on resource-constrained devices including mobile terminals.

Keywords: classification of wild mushrooms, ShuffleNetV2, deep learning, lightweight

DOI: [10.25165/j.ijabe.20251801.9179](https://doi.org/10.25165/j.ijabe.20251801.9179)

Citation: Xu X M, Yang D W, Wei Q Q, Li J Y, Zhang J. Classification of wild mushrooms based on improved ShuffleNetV2. *Int J Agric & Biol Eng*, 2025; 18(1): 208–218.

1 Introduction

China has been reported to possess abundant fungal resources, with 1662 known species of wild mushrooms. Notably, 1020 are edible, 692 are medicinal, and 480 are poisonous^[1-3]. Due to the complex growth environments of wild mushrooms and a lack of knowledge regarding mushroom identification, poisoning incidents frequently happen^[4,5]. Therefore, the accurate and rapid identification of wild mushrooms has become an urgent and crucial duty.

Advances in artificial intelligence and computer technology have enabled effective crop identification and disease diagnosis through computer vision^[6-10]. Deep learning trained on large-scale data has produced high-accuracy image recognition models that surpass traditional methods^[11-13]. Wang et al.^[14] developed a recognition model using support vector machines, random forests (RF), and convolutional neural networks (CNNs), trained on 350 images, with the RF model achieving the highest accuracy. Lin et al.^[15] researched the identification of wild mushrooms using

machine vision and image processing techniques, achieving a 90.78% accuracy rate with the proposed method by the Retinex algorithm for image enhancement. Although machine learning-based research models are less complex, their accuracy is lower^[16], whereas the models based on convolutional neural networks excel in image recognition tasks. Convolutional neural networks excel at extracting and learning data features, simplifying tasks while providing high accuracy, and automating feature extraction in image recognition^[17-20]. For example, Ketwongsa et al.^[21] proposed the poisonous mushroom classification using an enhanced AlexNet model with the initial module of GoogLeNet, achieving an accuracy of 95.5%. Peng et al.^[22] introduced a Multi-Dimensional Feature Fusion Attention Network, M-ViT, incorporating the MDA module, with M-ViT achieving 96.21% accuracy on mushroom datasets and 91.83% on the MO106 dataset. Despite the accuracy improvements offered by deep learning-based recognition models, they also escalate model complexity and hardware costs.

Extensive research has been conducted on lightweight models. Xiao et al.^[23] trained a ShuffleNetV2 model using 1675 mushroom images, achieving a Top-1 accuracy of 55.18% and a Top-5 accuracy of 93.55%. The relatively small size of the dataset contributes to the low Top-1 accuracy. Wan et al.^[24] proposed a novel lightweight CNN model featuring a multi-feature partitioning design to reduce model complexity. This model achieved a 94.6% accuracy on the CIFAR-100 dataset, while reducing the number of parameters. Zhu et al.^[25] introduced MobileNetV3_large for shiitake mushroom grading classification, enhancing the SE attention module, using PolyFocalLoss, and employing transfer learning. The improved model achieved 99.91% accuracy with a 26.54% reduction in model size. Maintaining high accuracy while

Received date: 2024-06-26 **Accepted date:** 2025-01-03

Biographies: Xingmei Xu, Professor, research interest: Fungal Phenotype, Email: xingmeix@jlau.edu.cn; Dawei Yang, Master candidate, research interest: Fungal Phenotype, Email: yangdawei@mails.jlau.edu.cn; Qiqi Wei, Master, research interest: computer vision, Email: qiqi.wei@qdc.edu.cn.

***Corresponding author:** Jinying Li, Professor, research interest: horticultural crop Science. College of horticulture, Jilin Agricultural University, Changchun 130118, China. Tel: +86-13578801427, E-mail: lijinying@jlau.edu.cn; Jian Zhang, Professor, research interest: smart agriculture. Faculty of Agronomy, Jilin Agricultural University, Changchun 130118, China. Department of Biology, University of British Columbia, Okanagan, Kelowna, BC V1V 1V7, Canada. Tel: +86-18801328549, E-mail: jian.zhang@ubc.ca.

lightweighting the model reduces computational and hardware costs, facilitating deployment on mobile devices. Research in wild mushroom classification and recognition primarily focuses on improving model accuracy by deepening the network structure^[26-28]. However, this approach, while effective in improving accuracy, overlooks the model’s lightweight design, resulting in computational overhead and deployment challenges in practical applications. Hence, investigating the highly accurate and lightweight wild mushroom classification model is crucial for its application and deployment. This study proposed an enhanced CHE_ShuffleNetV2 model that reduced the model parameters and computation complexity while maintaining classification accuracy. The study presented here aimed to achieve the following specific objectives: 1) the CSP structure was introduced, employing a cross-stage local connection method to fully utilize the extracted feature information, effectively reducing the number of model parameters and computational complexity; 2) the ordinary convolution in the base unit was replaced with dilated convolution to expand the

receptive field and enhance the feature extraction capability of the model. A mixed dilated convolution strategy was employed to mitigate the loss of partial detail information; 3) the ECA channel attention mechanism was added to effectively extract feature information from important regions in the image, thereby enhancing the model’s recognition capability; 4) the ReLU activation function was replaced with the Hardswish activation function to reduce the computational load.

2 Materials and methods

2.1 Data acquisition

The dataset used in this study comprised two parts. The first part consisted of a mushroom dataset published on Kaggle, and the second part comprised the images of wild mushrooms captured in Fusong, Tonghua, Jilin, and other locations, all featuring complex backgrounds. The dataset comprised 7017 images of ten species of wild mushrooms, including Agaricus. The samples of these wild mushrooms are shown in Figure 1.



Figure 1 Images of some wild mushroom samples

2.2 Data preprocessing and partitioning

The scale of a dataset can affect the recognition performance of the model. Data augmentation operations such as increasing/decreasing brightness, adding Gaussian noise, horizontal flipping,

rotating, translating, and enhancing contrast can mitigate overfitting and enhance the generalization ability of the model^[29,30]. This study employed these operations to enhance the dataset. The augmentation results are shown in Figure 2.

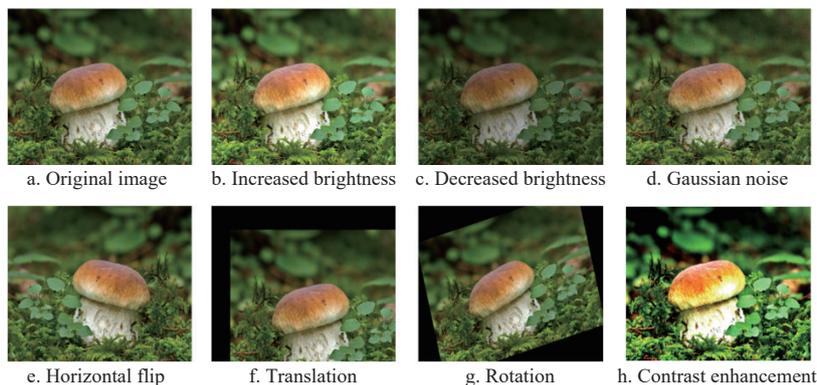


Figure 2 Images of data augmentation effects

After the data augmentation, the dataset contained 11 007 images. The dataset was divided into training, validation, and test sets at an 8:1:1 ratio. The information for each category of wild mushrooms is presented in Table 1.

2.3 Improved ShuffleNetV2 wild mushroom classification model

ShuffleNetV2, a lightweight network model introduced by Zhang et al.^[31], is based on ShuffleNetV1. The network structure is shown in Figure 3.

The ShuffleNetV2 network comprises base and down sampling

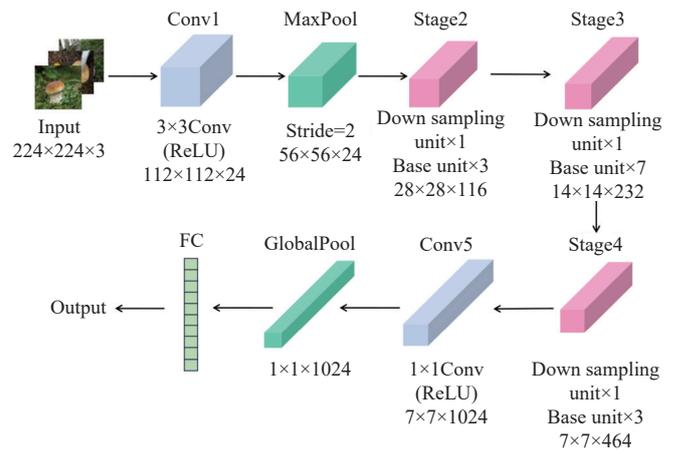
units, as shown in Figure 4. Figure 4a illustrates the base unit that partitions the input feature map into channels. The left branch performs identity mapping, and the right branch performs feature extraction using convolutional layers. Subsequently, the two branches are concatenated via Concat and undergo a Channel Shuffle. Importantly, the base unit maintains the feature-map dimensions and channel count.

Figure 4b depicts the down sampling unit, where both branches employ deep convolution with a stride of two for dimensionality reduction. The two branches are then concatenated via Concat and

subjected to a Channel Shuffle. This unit reduces the feature map size by half, while doubling the number of channels.

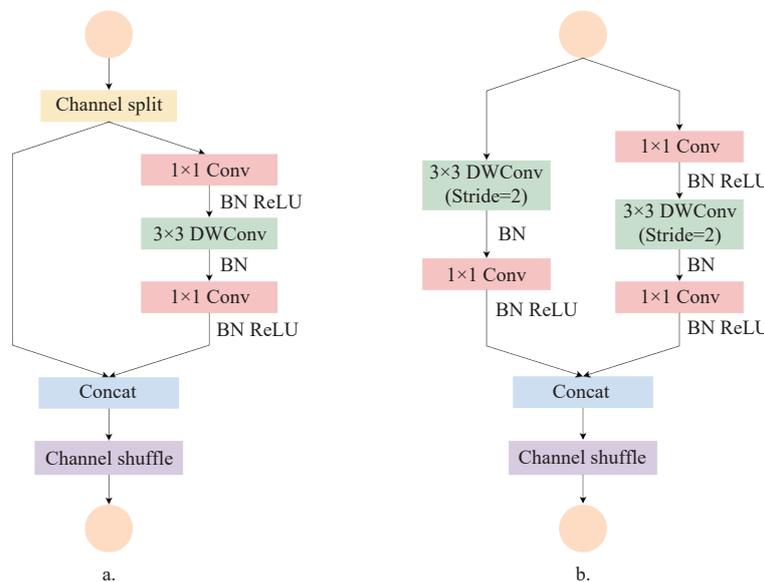
Table 1 Number of different types of wild mushrooms

Category	Before augmentation (Number of images)	After augmentation (Number of images)
Agaricus	347	1103
Amanita	741	1100
Boletus	1071	1071
Cantharellus	800	1060
Cortinarius	830	1114
Entoloma	362	1092
Hygrocybe	314	1107
Lactarius	1104	1104
Russula	1147	1147
Suillus	301	1109
Total	7017	11 007



Note: The stage module contains its down sampling units and base units, with FC representing fully connected layers.

Figure 3 Structure of ShuffleNetV2



Note: DWConv is depthwise convolution; a. base unit; b. down sampling unit; and the left branch of the down sampling unit adds a depthwise convolution.

Figure 4 Constituent units of ShuffleNetV2

ShuffleNetV2 strikes an effective balance between computational complexity and performance. By optimizing channel shuffle and depthwise convolutions, it significantly reduces computational overhead, which is critical for the wild mushroom classification task in this study.

2.3.1 CHE_ShuffleNetV2

ShuffleNetV2 has shown strong performance in various applications^[32,33]. Current research on wild mushroom classification mainly aims to improve accuracy by deepening network models. However, stacking convolutional modules increases model parameters and computational cost. To enhance classification accuracy and ensure efficient deployment on mobile devices, this study introduced an enhanced CHE_ShuffleNetV2 model based on ShuffleNetV2, integrating CSP, HDC, the ECA module, and the Hardswish activation function. The model expands the receptive field, enhances feature selection, and improves classification performance in complex environments.

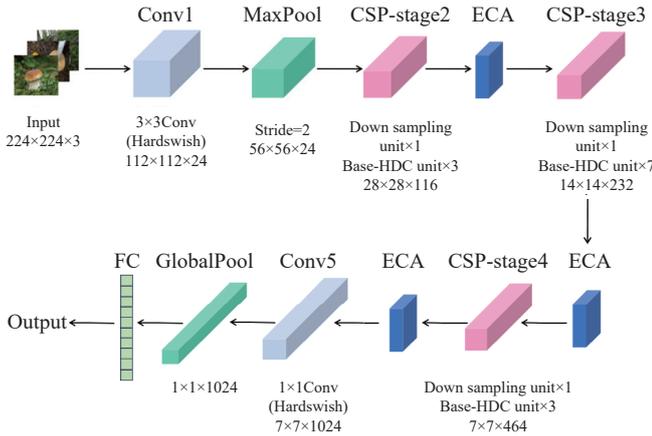
The CSP structure was integrated into the stage module (CSP-stage module), leveraging cross-stage local connections to enrich the gradient information, reduce gradient reuse, and reduce the computational load. By partitioning channels and sharing local

information, the model provides richer gradient information across different stages, enabling efficient learning of complex wild mushroom features. The three CSP stages consisted of down sampling units and base-HDC units, with the number of these units in the three module layers set at 1:3, 1:7, and 1:3, respectively. The Deepwise convolution in the base unit was substituted with Deepwise Dilated convolution to form the base-HDC unit. The Base-HDC unit employs a Hybrid Dilated Convolution structure, which eliminates the grid effect seen in conventional dilated convolutions, resulting in a more uniform receptive field expansion and improving the model’s ability to recognize image details. The ECA module was added after the CSP-stage layer to adaptively weight channel features, enabling the model to focus on key features for distinguishing different wild mushroom species. Additionally, the Hardswish activation function replaced the ReLU activation function to enhance the model performance. The overall structure of the CHE_ShuffleNetV2 model is illustrated in Figure 5.

2.3.2 Cross-stage partial

CSP^[34] optimizes gradient propagation and reduces computational complexity in neural networks. The principle involves dividing the input feature map’s channel dimensions into

two equal parts. One part is passed directly to the next convolutional layer, while the other undergoes convolution before merging with the first. This parallel processing and feature reorganization reduce redundant computations. By processing only a subset of channels with convolution and weighting after division, the CSP structure reduces parameters and computational load, optimizing efficiency while maintaining accuracy. CSP employs skip connections to fuse features from different layers, optimizing gradient flow balance and enhancing model stability in deep networks. This design reduces parameters and enhances the model's learning effectiveness.



Note: The gray part represents the convolution module with Hardswish, the pink part represents the CSP-stage module with the CSP structure, the blue part represents the ECA module, and the base-HDC unit represents the base unit with HDC.

Figure 5 Structure of CHE_ShuffleNetV2

In this study, a 1×1 convolution block is incorporated into the left branch of the CSP structure to facilitate cross-channel information interaction, enhancing the model's ability to capture inter-channel dependencies. This design allows the model to acquire rich information from the input feature map at a lower computational cost. The right branch employs the Base-HDC unit to further extract features from the input feature map. After feature extraction in both branches, the outputs are concatenated through cross-stage partial connections, generating a more diverse set of feature combinations. The output is further optimized through channel shuffling, enhancing channel information integration and improving the model's expressiveness. The enhanced CSP-stage structure is illustrated in Figure 6.

2.3.3 Hybrid Dilated Convolution

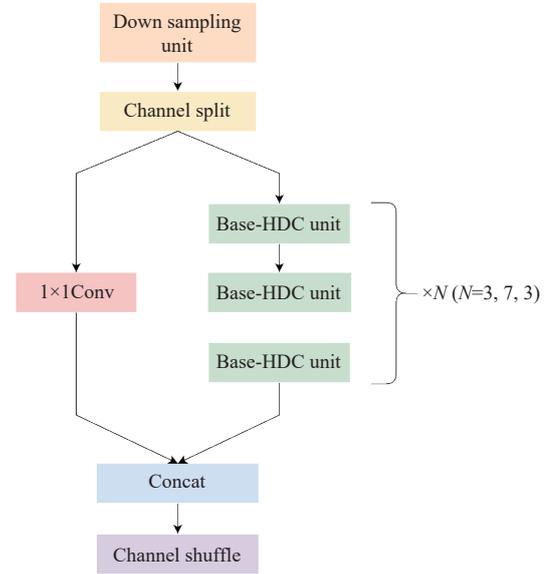
The dilated convolution^[35] extends the feature extraction range of the model by enlarging the receptive field of the convolution kernel while maintaining its size. Compared to the traditional convolution, the dilated convolution adjusts the receptive field size using the dilation rate, as shown in Equation (1).

$$k' = k + (k - 1) \times (r - 1) \quad (1)$$

where, k represents the size of the original convolution kernel, r represents the dilation rate, and k' represents the size of the receptive field. Figure 7 compares ordinary convolution with dilated convolution, where the dark blue area indicates the receptive field. In Figure 7a, the receptive field size of an ordinary convolution map is 3×3 , whereas in Figure 7b, the receptive field size of a dilated convolution map is 5×5 .

An excessively large receptive field can yield sparse feature information, leading to insufficiently extracted features in the

shallow layers^[36]. Stacking convolutional layers with the same dilation rate can induce a grid effect, leading to a loss of local key information, as shown in Figure 8. The use of three consecutive dilated convolutions with the same dilation rate may not extract all the pixels, resulting in the loss of partial detail information.



Note: The left branch is added to form a CSP structure with the right branch, and the base-HDC unit is the convolution unit with HDC added.

Figure 6 Structure of the CSP-stage module

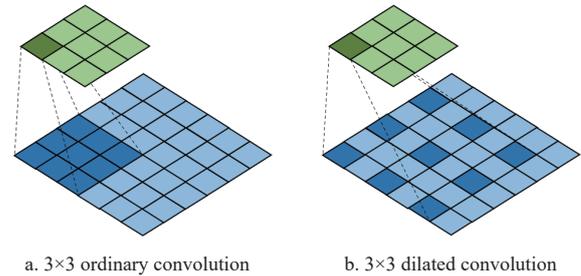


Figure 7 Ordinary convolution and dilated convolution

The HDC employs multiple dilated convolutions with varying dilation rates to extract comprehensive feature information^[37]. In this approach, each sample has a different dilation rate r , with the values of 1, 2, and 3, as illustrated in Figure 9a. The sampling results, as shown in Figure 9b, ensured the extraction of all pixels within the receptive field to address the issue of information loss. The dilation rate must satisfy Equation (2).

$$M_i = \max[M_{i+1} - 2r_i, 2r_i - M_{i+1}, r_i] \quad (2)$$

where, M_i represents the maximum distance between two non-zero elements in layer i , and r_i is the dilation rate of the i -th layer.

Replacing the Depthwise convolution in the right branch of the base unit with an HDC enlarges the receptive field and improves the feature extraction capability, as depicted in Figure 10.

2.3.4 ECA mechanism

The ECA mechanism is a lightweight channel attention method that boosts channel expressiveness without adding significant parameters or computations. The fundamental principle of ECA is capturing channel correlations through local interactions, enabling the model to focus on key features.

The ECA mechanism compresses the spatial features of the input feature map ($C \times H \times W$) into a global feature value via global

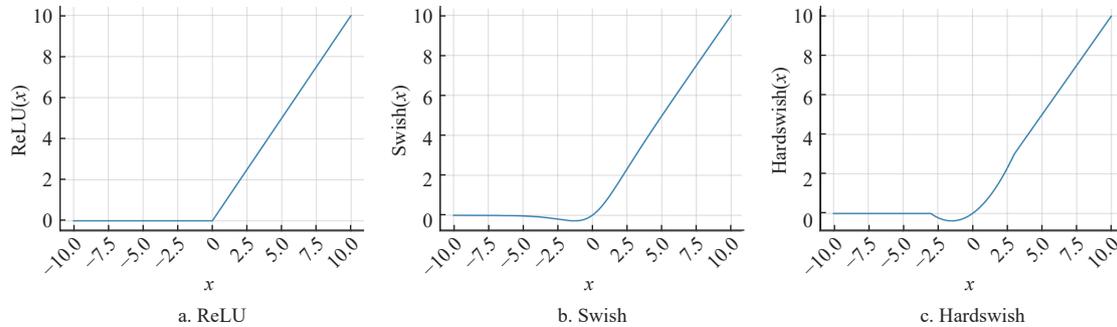
$$\text{Swish}(x) = x \cdot \sigma(x) \quad (5)$$

The Hardswish activation function builds upon enhancement of the Swish activation function. It not only addresses the issue of neuron death, but also employs a piecewise linear function to eliminate the exponential operation found in the Swish function, resulting in faster computation. Figure 12c illustrates the function image of the Hardswish activation function, with Equations (6) and

(7) specifying the calculation.

$$\text{ReLU6}(x) = \min(\max(x, 0), 6) \quad (6)$$

$$\text{Hardswish}(x) = x \cdot \frac{\text{ReLU6}(x+3)}{6} = \begin{cases} 0, & x \leq -3 \\ x, & x \geq +3 \\ x \cdot \frac{(x+3)}{6}, & \text{otherwise} \end{cases} \quad (7)$$



Note: The horizontal axis represents input feature maps, and the vertical axis represents output feature maps of the activation function.

Figure 12 Graphs of the activation functions

2.4 Experimental environment and settings

The network model used in this experiment was trained using the PyTorch deep-learning framework. The experimental hardware comprised an Intel(R) Xeon(R) Gold 6246R@3.40GHz CPU with 32 cores and 128GB of memory, and an NVIDIA Quadro RTX 8000 GPU with 48GB of video memory. All experiments were conducted using the Windows 10 operating system running Python version 3.8.

To enhance the convergence rate of the model and reduce the training time, the experiment utilized the pre-trained ShuffleNetV2_1x model. Details of the hyperparameter settings are listed in Table 2.

Table 2 Hyperparameter settings

Parameters	Value
Batch_size	32
Image size	224×224
Learning rate	0.001
Epoch	200
Optimizer	Adam
Loss function	Cross Entropy

2.5 Evaluation indices

In this study, Accuracy, Precision, Recall, and F1-Score were selected as the evaluation indices of model performance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

where, TP means the prediction is positive and the actual is positive; FP means the prediction is positive and the actual is negative; TN means the prediction is negative and the actual is negative; FN means the prediction is negative and the actual is positive.

To better evaluate the lightweight performance of the model, parameters and FLOPs operations were selected to assess the computational complexity of the model.

3 Results

3.1 Comparison of model performance of different attention mechanisms

Three common attention mechanisms, including CBAM^[40], SE^[41], and ECA, were selected for a comparative analysis of their impact on the performance of the model. Table 3 presents the results.

Table 3 Comparison of different attention mechanisms

Model	Acc/%	P/%	R/%	F1-Score/%
ShuffleNetV2	92.21	91.37	91.48	90.35
ShuffleNetV2+CBAM	91.10	90.82	90.60	89.20
ShuffleNetV2+SE	92.40	92.41	92.00	91.17
ShuffleNetV2+ECA	93.92	93.10	93.00	92.32

Note: CBAM refers to Convolutional Block Attention Module, SE refers to Squeeze-and-Excitation, and ECA refers to Efficient Channel Attention. Same as below.

Table 3 shows that each metric of ShuffleNetV2+CBAM decreased, whereas those of ShuffleNetV2+SE and ShuffleNetV2+ECA increased, with ShuffleNetV2+ECA yielding the highest scores. The accuracy, precision, recall, and F1-Score of ShuffleNetV2+ECA were 93.92%, 93.10%, 93.00%, and 92.32%, respectively, indicating improvements of 1.71%, 1.73%, 1.52%, and 1.97% over ShuffleNetV2.

3.2 Comparison of model performance of different activation functions

Three common activation functions, Swish, ELU, and Hardswish, were selected to analyze their impact on model performance. Table 4 presents the results.

Table 4 Comparison of different activation functions

Model	Acc/%	P/%	R/%	F1-Score/%	FLOPs/M
ShuffleNetV2(ReLU)	92.21	91.37	91.48	90.35	149.58
ShuffleNetV2+Swish	92.27	91.27	91.33	90.00	149.60
ShuffleNetV2+ELU	92.50	92.00	92.70	91.30	149.58
ShuffleNetV2+Hardswish	92.65	92.33	92.77	92.26	148.10

Table 4 reveals that using Swish as the activation function resulted in a model accuracy of 92.27%, a mere 0.06% improvement over that of ShuffleNetV2. However, the precision, recall, and F1-Score decreased by 0.10%, 0.15%, and 0.35%, respectively. The FLOPs increased slightly to 149.60 M by 0.02 M. Conversely, employing ELU and Hardswish as activation functions improved all the metrics, with Hardswish outperforming ELU. Specifically, ShuffleNetV2+Hardswish achieved accuracy, precision, recall, and F1-Score of 92.65%, 92.33%, 92.77%, and 92.26%, respectively, surpassing ShuffleNetV2 by 0.44%, 0.96%, 1.29%, and 1.91%, respectively. Additionally, the FLOPs decreased to 148.10 M, a reduction of 1.48 M.

3.3 Ablation experiment

Ablation experiments were conducted to investigate the effects of the CSP structure, HDC, ECA module, and Hardswish activation function on the ShuffleNetV2 model performance. Table 5 presents the results.

Table 5 shows that with the introduction of the CSP structure, the model achieved an accuracy, precision, recall, and F1-Score of 92.55%, 91.42%, 92.16%, and 90.70%, respectively, with 0.924 M parameters and 103.50 M FLOPs. Compared with ShuffleNetV2,

this represented a modest improvement of 0.34%, 0.05%, 0.68%, and 0.35%, respectively, while significantly reducing the parameters and FLOPs by 26.84% and 30.80%, respectively. Substituting common convolution with HDC resulted in an accuracy, precision, recall, and F1-Score of 92.93%, 92.45%, 91.67%, and 91.60%, respectively, with 1.282 M parameters and 153.21 M FLOPs. Compared with ShuffleNetV2, the model's accuracy, precision, recall, and F1-Score increased by 0.72%, 1.08%, 0.19%, and 1.25%, respectively, with only a slight increase in parameters (0.019 M) and FLOPs (3.63 M). Upon adding the ECA module, the model achieved accuracy, precision, recall, and F1-Score of 93.92%, 93.10%, 93.00%, and 92.32%, respectively, with 1.263 M parameters and 149.58 M FLOPs. Compared with ShuffleNetV2, while maintaining the number of parameters and FLOPs, these metrics increased by 1.71%, 1.73%, 1.52%, and 1.97%, respectively. Using the Hardswish activation function resulted in accuracy, precision, recall, and F1-Score of 92.65%, 92.33%, 92.77%, and 92.26%, respectively, with 1.263 M parameters and 148.10 M FLOPs. Compared to ShuffleNetV2, with unchanged parameters, these metrics improved by 0.44%, 0.96%, 1.29%, and 1.91%, respectively, whereas FLOPs decreased by 1.48 M.

Table 5 Ablation experiment

Model	CSP	HDC	ECA	Hardswish	Acc/%	P/%	R/%	F1-Score/%	Param/M	FLOPs (M)
	-	-	-	-	92.21	91.37	91.48	90.35	1.263	149.58
	√	-	-	-	92.55	91.42	92.16	90.70	0.924	103.50
	-	√	-	-	92.93	92.45	91.67	91.60	1.282	153.21
	-	-	√	-	93.92	93.10	93.00	92.32	1.263	149.58
	-	-	-	√	92.65	92.33	92.77	92.26	1.263	148.10
	√	√	-	-	93.44	93.31	92.63	91.95	0.933	105.32
	√	-	√	-	94.31	93.79	94.17	93.26	0.924	103.50
ShuffleNetV2	√	-	-	√	93.51	92.77	93.10	92.14	0.924	102.27
	-	√	√	-	94.80	94.63	94.32	93.60	1.282	153.21
	-	√	-	√	94.01	93.65	93.93	92.90	1.282	151.72
	-	-	√	√	94.34	93.94	94.19	93.20	1.263	148.10
	√	√	√	-	94.67	94.35	94.43	93.48	0.933	105.32
	√	√	-	√	94.21	93.00	93.17	92.18	0.933	104.08
	√	-	√	√	94.41	93.69	93.94	93.10	0.924	102.27
	-	√	√	√	94.65	94.39	94.43	93.48	1.282	151.72
CHE_ShuffleNetV2	√	√	√	√	95.02	95.19	94.56	94.00	0.933	104.08

The CHE_ShuffleNetV2 model achieved accuracy, precision, recall, and F1-Score of 95.02%, 95.19%, 94.56%, and 94.00%, respectively, representing improvements of 2.81%, 3.82%, 3.08%, and 3.65%, respectively, over the original model. Regarding model complexity, the CHE_ShuffleNetV2 model had 0.933 M parameters and 104.08 M FLOPs, marking reductions of 26.13% and 30.42%, respectively, compared to the original model.

3.4 Comparison of performance of different models

Figure 13 compares the training results of the lightweight convolutional neural network models, including ShuffleNetV2, EfficientNet^[42], DenseNet^[43], MobileNetV2^[44], and CHE_ShuffleNetV2. As shown in Figure 13a, the accuracy of each model steadily increased and converged with the number of iterations. Notably, CHE_ShuffleNetV2 consistently outperformed other models in terms of accuracy throughout the training. Figure 13b shows that the loss values of each model progressively decreased and converged with increasing iterations, with the loss curve of CHE_ShuffleNetV2 approaching 0.

The experimental results of the model for the test set are presented in Table 6.

Table 6 presents that EfficientNet exhibited the lowest accuracy, precision, recall, and F1-Scores at 90.51%, 88.66%,

88.90%, and 87.38%, respectively, with 4.020 M parameters, 398.03 M FLOPs, and a model size of 15.998 MB. DenseNet achieved accuracy, precision, recall, and F1-Score of 91.13%, 90.62%, 92.05%, and 90.97%, respectively, with the largest number of parameters (6.964 M), FLOPs (2880.00 M), and model size (27.793 MB). MobileNetV2 demonstrated the highest accuracy, precision, recall, and F1-Score at 92.32%, 92.24%, 92.05%, and 90.97%, respectively, with 2.236 M parameters, 318.97 M FLOPs, and a model size of 8.975 MB. ShuffleNetV2 achieved slightly lower metrics than MobileNetV2, with accuracy, precision, recall, and F1-Score of 92.21%, 91.37%, 91.48%, and 90.35%, respectively, and the smallest number of parameters (1.263 M), FLOPs (149.58 M), and model size (5.106 MB).

CHE_ShuffleNetV2 achieved an accuracy, precision, recall, and F1-Score of 95.02%, 95.19%, 94.56%, and 94.00%, respectively, with 0.933 M parameters, 104.08 M FLOPs, and a model size of 3.833 MB. Compared with MobileNetV2, it showed improvements of 2.70%, 2.95%, 2.51%, and 3.03% in accuracy, precision, recall, and F1-Score, respectively. Compared with ShuffleNetV2, CHE_ShuffleNetV2 reduced the number of parameters, FLOPs, and model size by 26.13%, 30.42%, and 25.00%, respectively.

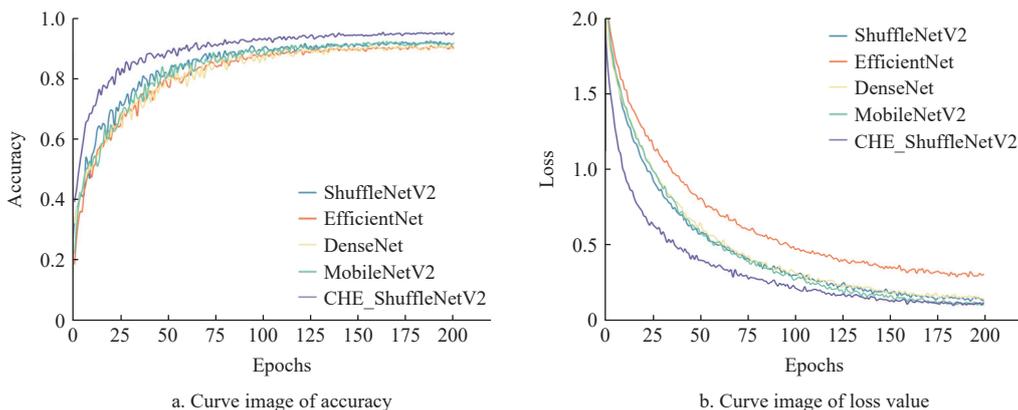


Figure 13 Accuracy of different model training and loss value change curves

Table 6 Test results of ShuffleNetV2, EfficientNet, DenseNet, MobileNetV2, and CHE_ShuffleNetV2

Model	Acc/%	P/%	R/%	F1-Score/%	Param/M	FLOPs/M	Model size/MB
ShuffleNetV2	92.21	91.37	91.48	90.35	1.263	149.58	5.106
EfficientNet	90.51	88.66	88.90	87.38	4.020	398.03	15.998
DenseNet	91.13	90.62	90.26	89.69	6.964	2880.00	27.793
MobileNetV2	92.32	92.24	92.05	90.97	2.236	318.97	8.975
CHE_ShuffleNetV2	95.02	95.19	94.56	94.00	0.933	104.08	3.833

Figure 14 shows the confusion matrix of the test results for ShuffleNetV2, EfficientNet, DenseNet, MobileNetV2, and CHE_ShuffleNetV2. The numbers on the main diagonal indicate the ratio of correct identification of each type of wild mushroom. CHE_ShuffleNetV2 exhibited smaller classification errors than the other models, achieving a 100% identification accuracy for *Agaricus*, *Entoloma*, *Hygrocybe*, and *Suillus*. *Amanita*, *Boletus*, and *Cantharellus* had over 95% identification accuracy, *Cortinarius* achieved 90% accuracy, and *Lactarius* and *Russula* had 87% accuracy. The discrepancies for these latter two types of wild

mushrooms could be attributed to the very subtle differences in their shape and texture, making them difficult to distinguish and resulting in classification errors.

3.5 Grad-CAM visualization

Grad-CAM (Gradient-weighted Class Activation Mapping) is a technique in deep learning for visualizing and understanding CNN decisions. It visualizes CNN decisions by highlighting key regions with heatmaps. Grad-CAM allows understanding of complex model decisions without sacrificing accuracy, striking a balance between interpretability and high performance, while preserving fidelity to the original model. In this study, Grad-CAM generates heatmaps of wild mushroom regions by computing gradients of predicted class scores from the final CNN convolutional layer’s activations. Grad-CAM analyzes gradients flowing into the final convolutional layer to highlight key regions of wild mushroom images. Grad-CAM enhances interpretability, aids in understanding CNN predictions, debugging, building trust, and revealing biases. As shown in Figure 15, compared to ShuffleNetV2, CHE_ShuffleNetV2 more effectively focuses on key regions of wild mushroom images in complex backgrounds^[45].

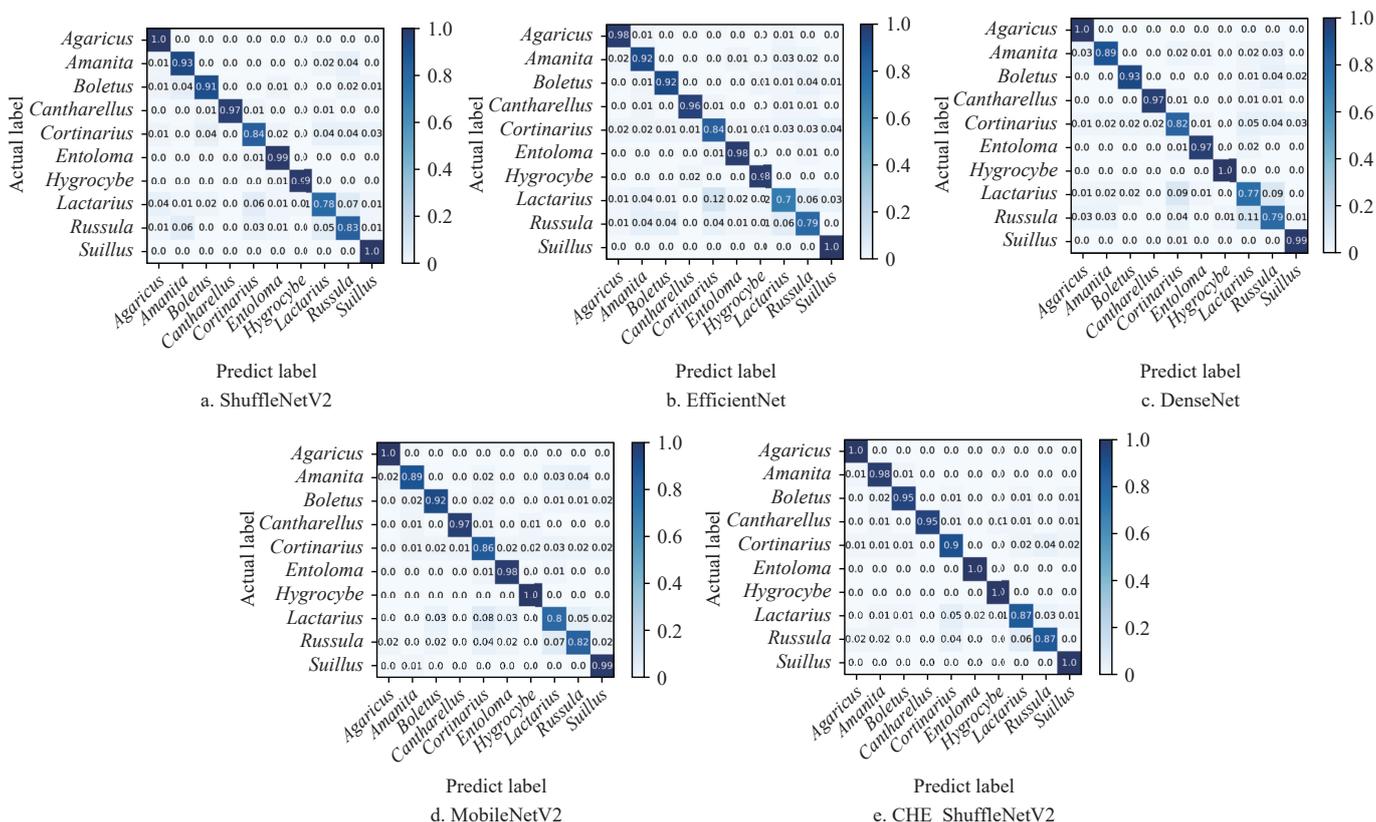


Figure 14 Confusion matrix

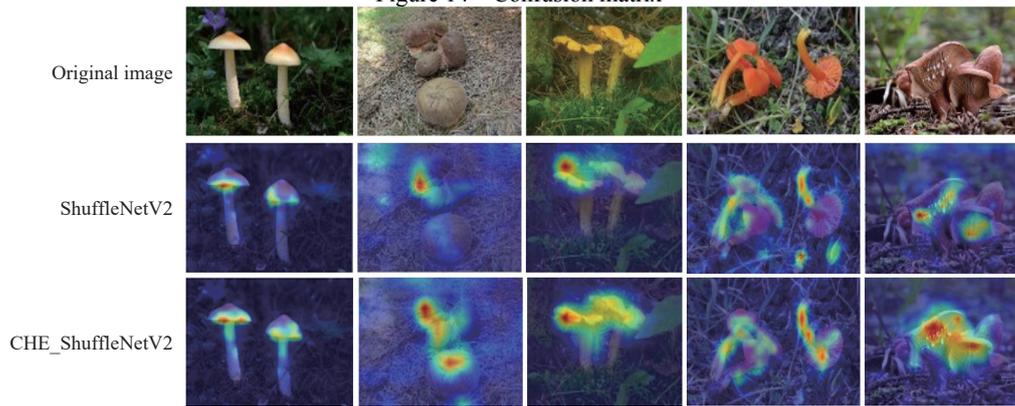


Figure 15 Visualization of the resulting thermodynamic diagram

4 Discussion

The image classification algorithm utilizing ShuffleNetV2 effectively categorized various wild mushroom species, which is comparable to prior studies^[46]. To enhance the accuracy of the model and decrease its complexity, this study conducted four primary tasks.

Utilizing a HDC, ShuffleNetV2+HDC achieved an accuracy, precision, recall, and F1-Score of 92.93%, 92.45%, 91.67%, and 91.60%, respectively. These values are 0.72%, 1.08%, 0.19%, and 1.25% higher than those of ShuffleNetV2. The increase in the number of parameters and FLOPs was minimal at only 0.019 M and 3.63 M, respectively. An HDC enhances the feature extraction capability of the model by expanding the receptive field^[47].

The addition of the ECA module resulted in accuracy, precision, recall, and F1-Score of 93.92%, 93.10%, 93.00%, and 92.32%, respectively, for ShuffleNetV2+ECA. This represents an increase of 1.71%, 1.73%, 1.52%, and 1.97%, respectively, compared with ShuffleNetV2, while keeping the number of parameters and FLOPs unchanged. The ECA module avoided the dimensionality reduction and utilized the 1-dimensional convolution for local cross-channel interactions, enhancing the extraction of channel dependencies^[48].

After introducing the CSP structure, ShuffleNetV2+CSP exhibited improvements in accuracy, precision, recall, and F1-Score. It reduced the number of parameters to 0.924 M and FLOPs to 103.50 M, reductions of 26.84% and 30.80%, respectively, compared to ShuffleNetV2. This indicated that the CSP structure effectively reduced redundant gradient information, thus lowering the network parameter count and computational requirements^[34].

Using the Hardswish activation function in ShuffleNetV2+Hardswish maintains the same number of parameters while improving the accuracy, precision, recall, and F1-Score. The FLOPs were reduced to 148.10 M compared to ShuffleNetV2, a decrease of 1.48 M. This reduction was due to the Hardswish activation function, which avoided the computation-intensive Swish function, thereby reducing the overall calculation load^[49].

The ablation experiments on interactions between module components - such as the CSP structure, HDC, ECA mechanism, and Hardswish activation - show that the CSP structure significantly reduces the number of parameters and floating-point operations. Adding HDC and the ECA mechanism significantly improves classification accuracy. Although adding HDC and the ECA mechanism increases parameters and floating-point operations, the increase is minimal. The CSP structure design effectively mitigates this issue. The model with the Hardswish activation function

effectively reduces floating-point operations without affecting classification accuracy.

The CHE_ShuffleNetV2 model achieved accuracy, precision, recall, and F1-Score of 95.02%, 95.19%, 94.56%, and 94.00%, respectively, representing increases of 2.81%, 3.82%, 3.08%, and 3.65% over ShuffleNetV2. It reduced the number of parameters to 0.933M and FLOPs to 104.08M, reductions of 26.13% and 30.42%, respectively, compared with ShuffleNetV2. This improved method enhanced the classification accuracy, reduced the model complexity, and was suitable for classifying wild mushrooms in complex environments.

The confusion matrix of CHE_ShuffleNetV2 shows an accuracy of 87% in recognizing the *Lactarius* and *Russula* species, slightly lower than for other species. This is due to the high similarity in shape and texture between these species, as well as their color and shape overlap with other objects in complex backgrounds, leading to misclassifications. This is linked to the model's ability to handle texture details and edge features, particularly in complex natural environments, where diverse backgrounds affect the model's judgment.

Grad-CAM visualization allows intuitive understanding of the model's focus on key regions of wild mushrooms during decision-making. The improved CHE_ShuffleNetV2 shows a significant increase in attention to key regions of wild mushrooms, focusing on larger areas while avoiding irrelevant regions like surrounding weeds.

Furthermore, the variations in climatic conditions (sunny, cloudy, and rainy) in images of wild bacteria also affected model training. The model was required to exhibit sufficient generalization to accurately classify bacteria under diverse conditions. In this study, artificial augmentation methods, such as blurring and brightness adjustment, were used to simulate wild mushroom images under varying climatic conditions. However, these augmented images still differ from real-world wild mushroom images to some extent. Future research will focus on expanding the dataset by collecting more wild mushroom images from diverse climatic conditions to enhance the model's generalization.

CHE_ShuffleNetV2 has enhanced the classification accuracy of complex wild mushroom images. Future research will explore the performance differences of the model on toxic and non-toxic wild mushrooms. Multi-scale feature fusion techniques will be used to extract and merge deep and shallow features, improving the model's ability to capture features for specific wild mushroom categories and enhancing accuracy in identifying toxic and non-toxic species.

5 Conclusion

This study addressed the challenge of low accuracy and complex structures of existing models by enhancing the lightweight network model, CHE_ShuffleNetV2. First, a dataset of wild mushrooms comprising 10 categories and 11 007 images was constructed. This study compared the effects of CBAM, SE, and ECA on model accuracy, as well as the effects of ReLU, Swish, ELU, and Hardswish on model complexity. The HDC replaced the common convolution, and the CSP structure was introduced. The results indicated that the ECA module most effectively improved the classification accuracy, while the Hardswish activation function significantly reduced the model complexity. The HDC enhanced the model accuracy, and the CSP structure reduced the complexity. The CHE_ShuffleNetV2 model achieved 95.02% accuracy with 0.933 M parameters and 104.08 M FLOPs. Compared with ShuffleNetV2, EfficientNet, DenseNet, and MobileNetV2, CHE_ShuffleNetV2 exhibited superior overall performance. Consequently, CHE_ShuffleNetV2 could accurately and rapidly classify wild bacterial species, making it suitable for deployment on resource-constrained devices, such as mobile terminals.

Acknowledgements

This work was supported by the Technology Development Plan Project of Jilin Province (Grant No. 20240304096SF), the Natural Science Foundation of Jilin Province (Grant No. YDZJ202201ZYTS544), and Jilin Agricultural University high-level researcher grant (Grant No. JLAUHLRG20102006). Data supporting the findings of this study are available from the corresponding author.

[References]

- [1] Li C T, Xu S. Edible mushroom industry in China: Current state and perspectives. *Applied microbiology and biotechnology*, 2022; 106(11): 3949–3955.
- [2] Wu F, Zhou L W, Yang Z L, Bau T, Li T H, Dai Y C. Resource diversity of Chinese macrofungi: edible, medicinal and poisonous species. *Fungal diversity*, 2019; 98: 1–76.
- [3] Zhang X R, Zhao H B, Zhang Y T, Wu K Y. Current situation, challenges, and future development directions of edible fungi industry in China. *Horticulture & Seed*, 2023; 43(5): 49–54, 97. (in Chinese)
- [4] Hamza A, Mylarapu A, Krishna K V, Kumar D S. An insight into the nutritional and medicinal value of edible mushrooms: A natural treasury for human health. *Journal of Biotechnology*, 2024; 381: 86–99.
- [5] Hou X L, Luo C Q, Chen S S, Zhang X W, Jiang J M, Yang Z F, et al. Progress in research on diseases of edible fungi and their detection methods: A review. *Crop Protection*, 2023; 174: 106420.
- [6] Mendoza-Bernal J, González-Vidal A, Skarmeta A F. A convolutional neural network approach for image-based anomaly detection in smart agriculture. *Expert Systems with Applications*, 2024; 247: 123210.
- [7] Qin C Y, Yang Y S, Gu F W, Chen P Y, Qin W C. Application and development of computer vision technology in modern agriculture. *Journal of Chinese Agricultural Mechanization*, 2023; 44(12): 119–128. (in Chinese)
- [8] Kaldarova M, Akanova A, Nazyrova A, Mukanova A, Tynykulova A. Identification of weeds in fields based on computer vision technology. *Eastern-European Journal of Enterprise Technologies*, 2023; 124(2): 44–52.
- [9] Luan X D. Research on the application of artificial intelligence and computer technology in agricultural modernization. *Advances in Computer, Signals and Systems*, 2022; 6(2): 6–10.
- [10] Tian L, Wang C, Li H L, Sun H T. Recognition method of corn and rice crop growth state based on computer image processing technology. *Journal of Food Quality*, 2022; 2022(1): 2844757.
- [11] Morales A, Villalobos F J. Using machine learning for crop yield prediction in the past or the future. *Frontiers in Plant Science*, 2023; 14: 1128388.
- [12] Thakur P S, Khanna P, Sheorey T, Ojha A. Trends in vision-based machine learning techniques for plant disease identification: A systematic review. *Expert Systems with Applications*, 2022; 208: 118117.
- [13] Yin H, Yi W L, Hu D M. Computer vision and machine learning applied in the mushroom industry: A critical review. *Computers and Electronics in Agriculture*, 2022; 198: 107015.
- [14] Wang Z J, Zhang L L, Li S F, Zhang, Q Y, Fu Q Q, Liu S J. Research on classification and recognition of edible fungi based on machine learning algorithm. *Journal of Fuyang Normal University (Natural Science)*, 2021; 38(4): 42–48. (in Chinese)
- [15] Lin N, Wang N, Li Z S, Chen X X, Zhang B, Tian X. Research on feature extraction and recognition of wild edible fungi based on machine vision. *Journal of Chinese Agricultural Mechanization*, 2020; 41(5): 111–119. (in Chinese)
- [16] Pradipkumar V H, Alagu Raja R A. Automatic identification of tree species from UAV images using machine learning approaches. *Journal of the Indian Society of Remote Sensing*, 2022; 50: 2447–2464.
- [17] Chen D G, Guli A, Yin P B, Lu Y N, Li S P. Identification of wild bacteria species based on improved Xception transfer learning. *Laser & Optoelectronics Progress*, 2021; 58(8): 245–254.
- [18] Espinosa-Herrera J M, Macedo-Cruz A, Fernández-Reynoso D S, Flores-Magdaleno H, Fernández-Ordóñez Y M, Soria-Ruiz J. Monitoring and identification of agricultural crops through multitemporal analysis of optical images and machine learning algorithms. *Sensors*, 2022; 22(16): 6106.
- [19] Ma X, Li Y L, Wan L P C, Xu Z X, Song J N, Huang J Q. Classification of seed corn ears based on custom lightweight convolutional neural network and improved training strategies. *Engineering Applications of Artificial Intelligence*, 2023; 120: 105936.
- [20] Xu P, Tan Q, Zhang Y P, Zha X T, Yang S M, Yang R B. Research on maize seed classification and recognition based on machine vision and deep learning. *Agriculture*, 2022; 12(2): 232.
- [21] Ketwongsa W, Boonlue S, Kokaew U. A new deep learning model for the classification of poisonous and edible mushrooms based on improved AlexNet convolutional neural network. *Applied Sciences*, 2022; 12(7): 3409.
- [22] Peng Y J, Xu Y, Shi J, Jiang S Y. Wild mushroom classification based on improved mobilevit deep learning. *Applied Sciences*, 2023; 13(8): 4680.
- [23] Xiao J W, Zhao C B, Li X J, Liu Z Y, Pang B, Yang Y H, et al. Research on mushroom image classification based on deep learning. *Software Engineering*, 2020; 23(7): 21–26. (in Chinese)
- [24] Wan G Q, Yao L. LMFRNet: A lightweight convolutional neural network model for image analysis. *Electronics*, 2023; 13(1): 129.
- [25] Zhu F W, Sun Y, Zhang Y Q, Zhang W J, Qi J. An improved MobileNetV3 mushroom quality classification model using images with complex backgrounds. *Agronomy*, 2023; 13(12): 2924.
- [26] Azadnia R, Noei-Khodabadi F, Moloudzadeh A, Jahanbakhshi A, Omid M. Medicinal and poisonous plants classification from visual characteristics of leaves using computer vision and deep neural networks. *Ecological Informatics*, 2024; 82: 102683.
- [27] Liu H, Hu Q R, Huang D Y. Research on the wild mushroom recognition method based on transformer and the multi-scale feature fusion compact bilinear neural network. *Agriculture*, 2024; 14(9): 1618.
- [28] Zhang F Y, Yin J L, Wu N, Hu X Y, Sun S K, Wang Y B. A dual-path model merging CNN and RNN with attention mechanism for crop classification. *European Journal of Agronomy*, 2024; 159: 127273.
- [29] Iqbal M J, Aasem M, Ahmad I, Alassafi M O, Bakhsh S T, Noreen N, et al. On application of lightweight models for rice variety classification and their potential in edge computing. *Foods*, 2023; 12(21): 3993.
- [30] Liang J, Jiang W P. A ResNet50-DPA model for tomato leaf disease identification. *Frontiers in Plant Science*, 2023; 14: 1258658.
- [31] Zhang X Y, Zhou X Y, Lin M X, Sun J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, Utah, USA: IEEE, 2018; pp.6848–6856.
- [32] Sun W H, Li Y, Feng H L, Weng X, Ruan Y P, Fang K, et al. Lightweight and accurate aphid detection model based on an improved deep-learning network. *Ecological Informatics*, 2024; 83: 102794.
- [33] Jin X, Jiang J X, Li Y, Wang Z. Improved ShuffleNetV2 for action recognition in BPPV treatment. *Biomedical Signal Processing and Control*, 2024; 88: 105601.

- [34] Wang C Y, Liao H Y M, Wu Y H, Chen P Y, Hsieh J W, Yeh I H. CSPNet: A new backbone that can enhance learning capability of CNN. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA: IEEE, 2020; pp.390–391.
- [35] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. arXiv: 1511.07122, 2015; In Press. doi: [10.48550/arXiv.1511.07122](https://doi.org/10.48550/arXiv.1511.07122).
- [36] Zhang X, Gao H Y, Wan L. Classification of fine-grained crop disease by dilated convolution and improved channel attention module. *Agriculture*, 2022; 12(10): 1727.
- [37] Wang P Q, Chen P F, Yuan Y, Liu D, Huang Z H, Hou X D, et al. Understanding convolution for semantic segmentation. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA: IEEE, 2018; pp.1451–1460. doi: [10.1109/WACV.2018.00163](https://doi.org/10.1109/WACV.2018.00163).
- [38] Wang Q L, Wu B G, Zhu P F, Li P H, Zuo W M, Hu Q H. ECA-Net: Efficient channel attention for deep convolutional neural networks. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA: IEEE, 2020; pp.11534–11542. doi: [10.1109/CVPR42600.2020.01155](https://doi.org/10.1109/CVPR42600.2020.01155).
- [39] Yang L, Xu S, Yu X Y, Long H B, Zhang H H, Zhu Y W. A new model based on improved VGG16 for corn weed identification. *Frontiers in Plant Science*, 2023; 14: 1205151.
- [40] Woo S, Park J, Lee J Y, Kweon I S. CBAM: Convolutional block attention module. In: Computer Vision - ECCV 2018, 2018; pp.3–19. doi: [10.1007/978-3-030-01234-2_1](https://doi.org/10.1007/978-3-030-01234-2_1).
- [41] Hu J, Shen L, Albanie S, Sun G, Wu E H. Squeeze-and-excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020; 42(8): 2011–2023.
- [42] Tan M X, Le Q V. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv: 1905.11946, 2019; In press. doi: [10.48550/arXiv.1905.11946](https://doi.org/10.48550/arXiv.1905.11946).
- [43] Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017; pp.2261–2269. doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).
- [44] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA: IEEE, 2018; pp.4510–4520. doi: [10.1109/CVPR.2018.00474](https://doi.org/10.1109/CVPR.2018.00474).
- [45] Jawadul K M, Faruq G M O, Md N, Mominul A, Julfikar H, Marcin K. Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad-CAM. *Scientific Reports*, 2024; 14: 16022.
- [46] Zhou H M, Chen J G, Niu X L, Dai Z G, Qin L, Ma L S, et al. Identification of leaf diseases in field crops based on improved ShuffleNetV2. *Frontiers in Plant Science*, 2024; 15: 1342123.
- [47] Li X C, Li X H, Zhang S M, Zhang G Y, Zhang M Q, Shang H Y. SLViT: Shuffle-convolution-based lightweight Vision transformer for effective diagnosis of sugarcane leaf diseases. *Journal of King Saud University-Computer and Information Sciences*, 2023; 35(6): 101401.
- [48] Ni H J, Shi Z W, Karungaru S, Lv S, Li X Y, Wang X X, et al. Classification of typical pests and diseases of rice based on the ECA attention mechanism. *Agriculture*, 2023; 13(5): 1066.
- [49] Li N, Xue J M, Wu S B, Qin K D, Liu N. Research on coal and gangue recognition model based on CAM-hardswish with efficientNetV2. *Applied Sciences*, 2023; 13(15): 8887.