

Comparison of Deep Transfer Learning Models for Pest Image Classification in Agriculture

Qixuan Huang *

Faculty of Computer Science & Information Technology, University Malaya, Kuala Lumpur, 50603, Malaysia

* Corresponding Author Email: 22065325@siswa.um.edu.my

Abstract. In the field of agriculture, crops are susceptible to attacks by pests. Accurately identifying and classifying crop pests, especially in their early stages of growth, is a significant challenge. Convolutional neural networks have become effective instruments for agricultural pest classification because of their capacity to extract and learn intricate information from photos. This study explores the use of transfer learning methods to compare efficient models with complex models to improve the effectiveness of pest and disease classification. The purpose of this study is to compare efficient models with complicated models using transfer learning techniques in order to increase the accuracy of pest and disease classification. By conducting experiments on the same pest and disease dataset, this research compares their performance in terms of accuracy, model size, computational resource consumption, and practical feasibility in real-world applications. The research results clearly demonstrate the performance comparison between different models, with efficient models excelling in some aspects while large models have advantages in others. This study contributes to better understanding for decision-makers in the agricultural sector on how to choose the appropriate deep learning models to enhance crop protection and increase yields.

Keywords: Convolutional neural network; transfer learning; pest classification.

1. Introduction

Crop pest detection and classification is a major difficulty in the realm of agriculture. Insects represent a major threat to crop production, and there is an urgent need to find effective pest management methods. However, accurately classifying pests is a challenging task due to the complex structures of insects and their high visual similarity across different species. Early identification and classification of pests in agricultural fields are particularly crucial, as this can help prevent pest infestations and assist in the selection of appropriate pesticides and biological control methods for managing crop diseases. Traditional manual pest identification methods often require a substantial amount of human labor, are time-consuming, and frequently exhibit low efficiency. To address these challenges in the field of agricultural research, deep learning technology has emerged as a critical tool.

When it comes to the identification and classification of crop pests, the significance of deep learning technology cannot be overstated. Traditional methods of insect classification typically rely on expert knowledge and involve a significant amount of time for classification. What sets deep learning algorithms apart is their ability to automatically learn complex features and patterns without the need for manual rule specification. By training on large-scale datasets, deep learning models can automatically extract features of pests and identify subtle differences, resulting in highly accurate classification. However, in practical applications, deep learning models often require a substantial amount of labeled data and computational resources, which may pose challenges in the agricultural domain.

One effective strategy to address these issues is transfer learning, which has gained popularity. Transfer learning lessens the need for large amounts of labeled data by enabling the information and features acquired from one activity to be applied to another related task. In the context of crop pest classification, transfer learning can help improve model performance, especially in scenarios with limited data or constrained resources. Through transfer learning, this study can leverage deep learning

models trained on other image classification tasks and fine-tune them for the specific task of pest and disease classification, thereby saving time and resources.

Therefore, both deep learning and transfer learning play crucial roles in addressing pest-related challenges in agriculture. This study aims to investigate transfer learning methods and compare the performance of efficient models such as MobileNet [1] and EfficientNet [2] with complex models like ResNet [3] and Xception [4] in the task of crop pest classification. This research poses the following research questions and research objectives: (1) how do efficient models and complex models perform in the classification of agricultural pests and diseases? (2) Can transfer learning enhance the performance of crop pest and disease classification?

The objective of this work is two-fold. The first one is to compare the accuracy of efficient models and complex models in the classification of agricultural pests, determining which model performs best in this task. The second one is to explore how transfer learning methods can improve the performance of crop pest classification and analyze their contributions to the models.

By achieving these research objectives, this research aims to provide valuable insights to decision-makers in the agricultural sector on how to select appropriate deep learning models to enhance crop protection and increase yields. It is believed that this study will contribute to driving technological innovation in the agricultural field and improving the sustainability of agricultural production.

2. Related work

In recent years, deep learning models based on convolutional neural network (CNN) have found extensive applications in image classification problems within the agricultural domain, encompassing areas such as plant disease recognition, fruit categorization, weed detection, and crop pest classification.

Wang et al. employed CNN for the classification of crop pest images. Building upon experiments involving the classification of pest images using LeNet-5 and AlexNet, they reconfigured the structure of the convolutional neural network for crop pest classification. They achieved a classification accuracy of 91% for 82 common pest types [5].

Lu et al. introduced a rice disease recognition method using deep Convolutional Neural Network (DCNN) technology, capable of accurately identifying ten common types of rice diseases, thereby improving both speed and accuracy of identification. They further incorporated transfer learning by fine-tuning pre-trained deep networks to enhance the learning efficiency [6].

C. Ngugi et al. conducted a comprehensive evaluation of recent research in plant disease detection using image processing techniques. They highlighted the replacement of shallow classifiers with deep learning algorithms, particularly when a substantial amount of training data is available. Techniques such as data augmentation and transfer learning were identified as methods to enhance classification accuracy [7].

Faithpraise Fina et al. employed k-means clustering and correspondence filtering for clustering and pest detection. The primary objective of this method is to identify and locate pests. This approach allows for the detection and identification of plant pests of various shapes, sizes, locations, and orientations using computationally efficient algorithms [8].

Thenmozhi Kasinathan et al. employed machine learning and insect pest identification methods for the classification and recognition of various insect datasets. They expanded their dataset through rescaling, preprocessing, and augmentation of insect photos. CNN models achieved accuracies of 91.5% for the Wang dataset and 90% for the Xie dataset [9].

Li. et al. developed a transfer learning-based residual network model for pest classification, achieving a notable accuracy of 86.95%, particularly when dealing with smaller datasets [10].

Mei et al. utilized image augmentation and transfer learning techniques for classifying different insect datasets. Among the various models tested, the VGG16 model exhibited the highest accuracy at 94.95%. They also recommended experimenting with different optimization methods and pre-trained models for further performance improvements [11].

In summary, CNN and transfer learning have played crucial roles in addressing image classification challenges within the agricultural domain, resulting in enhanced classification accuracy and performance.

3. Methodology

3.1. Dataset

In this study, an open-source insect dataset is leveraged, comprising samples of 12 common crop pests, as shown in Fig 1. The dataset consists of 3516 training samples and 1099 testing samples, providing us with a rich source of data for experiments [12]. The distribution of images in the dataset is shown in the Fig 2.

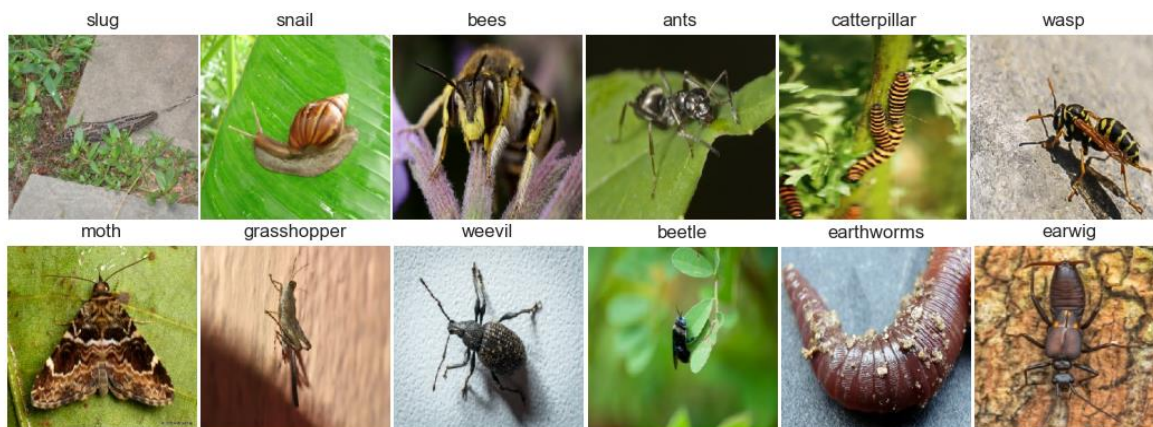


Fig 1. Examples of insect classes in the dataset [12].

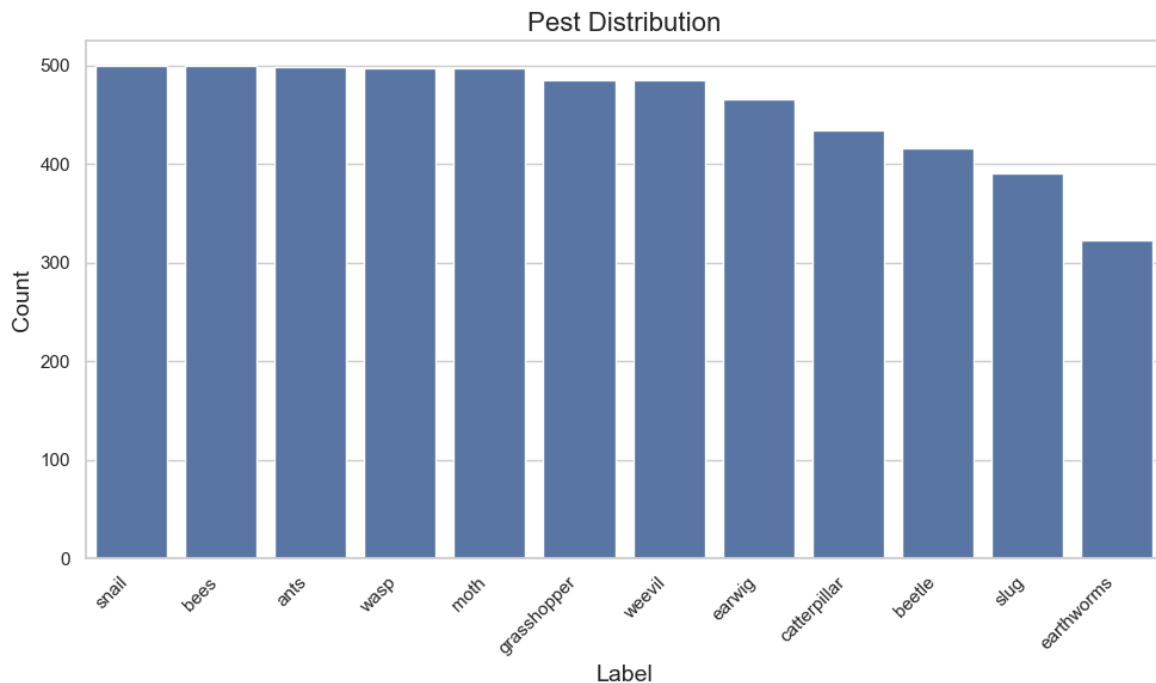


Fig 2. Pest distribution in insect dataset (Figure Credits: Original).

The diversity and quantity of these data samples are crucial for training and evaluating deep learning models. Data collection and preparation are essential steps in this research. This research applied data augmentation techniques, such as scaling, flipping, rotation, and transposition, to enhance the diversity of the dataset, better aligning it with the classification task. This approach not only increased the quantity of training data but also contributed to preventing model overfitting, thereby enhancing the performance and generalization ability of the classification model.

3.2. Convolutional Neural Network Preliminary

Convolutional neural networks, which perform exceptionally well in picture classification, are among the most widely used models in the agricultural field. Convolutional, pooling, and fully connected layers make up CNN. Convolutional layers utilize pooling layers to prevent overfitting and extract characteristics from insect pictures. The last completely linked layer uses a softmax activation algorithm to classify photos of insects.

However, complex CNN models are prone to overfitting, especially with limited training data, and require additional computational resources and time due to their numerous parameters. These challenges must be addressed in the context of pest classification.

3.3. Models

In this study, transfer learning techniques are employed by leveraging pretrained models that have accumulated rich knowledge through training on large-scale datasets. These models excel at capturing various generic features such as textures, shapes, and edges, providing a robust foundation for specific tasks. This approach of transfer learning typically significantly reduces the amount of data and time required for model training.

However, for the specific task of insect classification, four prominent models were selected (ResNet, EfficientNet, MobileNet, and Xception) as feature extractors. These four models have earned excellent reputations in the field of computer vision, each possessing its unique strengths and limitations.

ResNet is a deep convolutional neural network known for its notable feature, the introduction of residual blocks. These residual blocks allow the network to go deeper, effectively addressing the gradient vanishing problem and aiding in learning more features and abstract representations. Due to its outstanding performance, ResNet finds widespread application in tasks like image classification [3].

EfficientNet is an efficient convolutional neural network architecture known for enhancing performance by optimizing model depth, width, and resolution. It employs a compound scaling method to achieve better performance under the same computational resources. Consequently, EfficientNet is often the preferred choice in resource-constrained environments [2].

MobileNet is a lightweight convolutional neural network specially designed for execution on mobile devices. It makes use of depth-wise separable convolution layers, which lower the number of parameters and computational complexity without sacrificing performance. This makes MobileNet an ideal choice for real-time image processing on resource-limited devices [1].

Xception is a deep convolutional neural network architecture known for its innovative approach to feature extraction. It introduces depth-wise separable convolutions, which enable it to capture intricate patterns and relationships within the data efficiently. This unique design leads to improved model performance and has made Xception particularly effective in tasks requiring high-level image feature extraction, such as image classification and object recognition [4].

Subsequently, fully connected layers were created for each model for the final classification task. This fully connected layer consists of 128 and 256 neurons, aiding in capturing higher-level features and associations in insect images. To address overfitting, this work introduced dropout layers after each fully connected layer, randomly deactivating some neurons to enhance the model's generalization ability. Using Resnet as an example, the model architecture is illustrated in Fig 3.

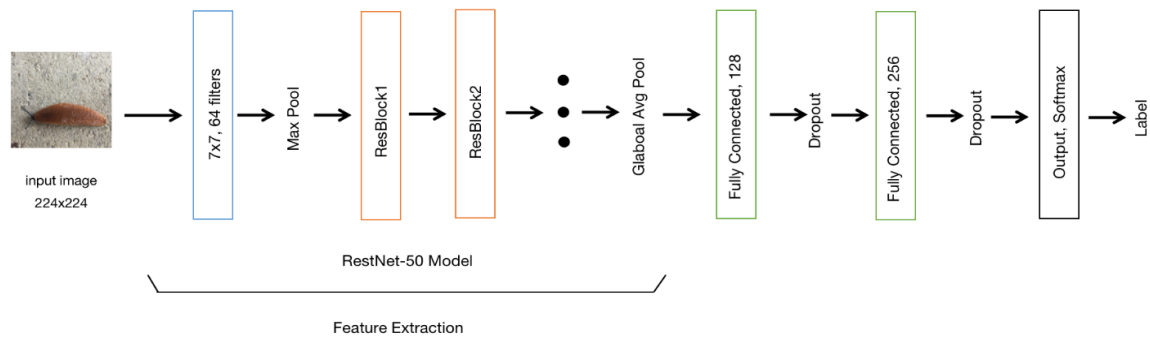


Fig 3. Model architecture (Figure Credits: Original).

The following sections will perform a comprehensive comparison of these four models to determine which one performs best in the insect classification task. This work will evaluate their performance, computational resource requirements, and other key factors to provide robust support and guidance for this research.

In summary, the core objective of this study is to identify the best-performing model among ResNet, EfficientNet, MobileNet, and Xception for the task of insect classification. Through this comparison, this work aims to gain deeper insights and essential decision-making guidance for the research.

3.4. Evaluation Metrics

In this section, the evaluation metrics are introduced to assess the performance of four classification models. The selection of appropriate evaluation metrics is crucial to gauge the effectiveness of the model in solving the problem at hand.

The following evaluation metrics were chosen based on their relevance to the research objectives, along with some additional performance metrics:

F1 Score: The F1 score is the harmonic mean of precision and recall and is used to measure the balance between precision and recall when classifying samples.

Accuracy: Accuracy represents the ratio of correctly classified samples to the total number of samples and serves as an indicator of the model's overall performance.

Recall: Recall quantifies the model's ability to correctly identify positive class samples, emphasizing the model's comprehensiveness.

Precision: Precision evaluates the accuracy of the model's positive class predictions by assessing the proportion of true positive samples among those predicted as positive.

Training Time: Training time refers to the amount of time, measured in seconds, required for the model to train on the training data.

Number of Parameters: The number of parameters reflects the complexity of the model and represents the quantity of parameters that the model needs to learn.

These evaluation metrics collectively provide a comprehensive assessment of classification model's performance, aligning with research objectives.

4. Result

4.1. Training details

All four models were trained using the same settings. The dataset comprises 4395 training samples and 1099 testing samples, each with dimensions of 224x224 pixels and 3 channels, distributed among 12 distinct classes. To facilitate training and validation, 20% of the training dataset was allocated as a validation set to monitor the models' performance.

Weight initialization involved deriving the model's weights from a pre-trained model, while the fully connected layer's weights were initialized with random values. To further reduce overfitting, a dropout rate of 0.5 was added to the fully connected layer. Model performance was measured during

training using categorical cross-entropy as the loss function and the Adam optimizer with a fixed learning rate of 0.0001. A total of 50 training iterations were performed on a MacBook Air with an M2 chip, and the training process was consistent across all four models.

Accuracy served as the primary performance metric, and early stopping was implemented to halt training if accuracy failed to improve or declined over a defined period.

4.2. Performances

In this section, this work will present the final output results of the four models, including: F1 score, Accuracy, Recall, and Precision, as shown in Table 1.

Table 1. Performance comparison.

Model	Accuracy	F1 score	Recall	Precision
ResNet50	0.8125	0.8014	0.8038	0.8121
Xception	0.8971	0.8899	0.8896	0.8948
EfficientNet B0	0.8826	0.8697	0.8695	0.8762
MobileNet	0.8598	0.8531	0.8530	0.8571

Through a performance comparison analysis, this study has reached the following conclusions: among the models under investigation, Xception exhibits the best performance, followed by EfficientNetB0, while ResNet50 demonstrates the poorest performance.

Next, this paper will delve further into the parameters and training time of these models. It's worth noting that the training duration is calculated based on the use of an early stopping strategy. Below is Table 2 providing relevant information on these aspects.

Table 2. Efficiency comparison.

Model	Params	Training time	Epochs	Model
ResNet50	23886092	20680s	47	ResNet50
Xception	21159860	22000s	11	Xception
EfficientNetB0	4249647	3168s	22	EfficientNetB0
MobileNet	3396172	3840s	32	MobileNet

Through the comparison of the two tables, a comprehensive conclusion could be draw: there is a significant disparity in training time and parameter count between complex models and high-performance models.

Specifically, the best performing model is Xception, requiring only 11 training epochs to achieve peak performance, but each training epoch takes around 2000 seconds. Following closely is EfficientNet, with a training time of approximately 160 seconds per epoch and a total training time of 3168 seconds. Its final accuracy is only 1.5% lower than Xception.

The validation loss and accuracy of four different models during training as a function of epochs are following in Fig 4 and Fig 5.

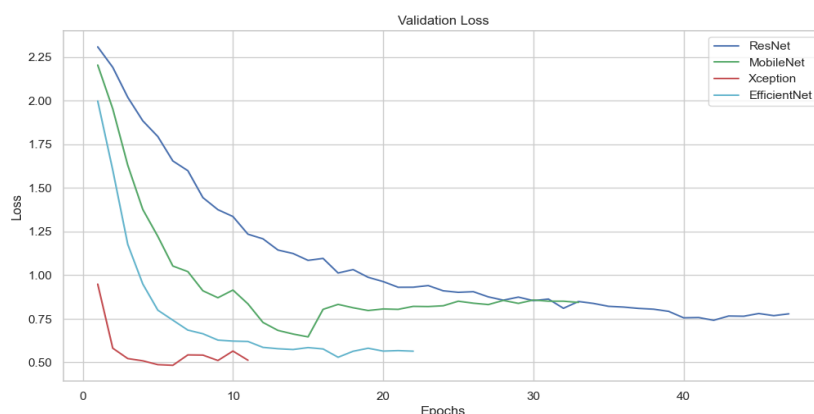


Fig 4. Comparison of validation loss over epochs (Figure Credits: Original).

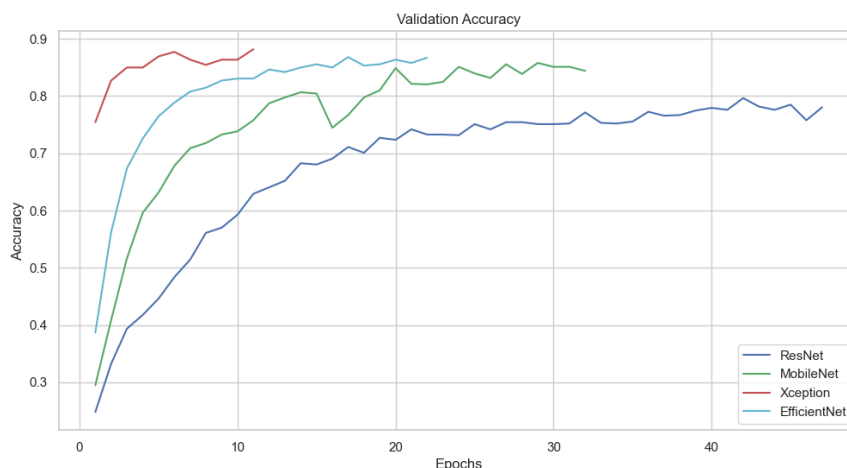


Fig 5. Comparison of validation accuracy over epochs (Figure Credits: Original).

Xception trains the fastest and already achieves 75% accuracy at the beginning. Next is EfficientNet, followed by MobileNet and ResNet.

5. Discussion

Using pre-trained models for pest classification has demonstrated excellent performance. Among the comprehensive evaluation of model size, training time, and accuracy etc., EfficientNet has performed the best, followed by Xception.

In terms of model design, the usage of Dropout in the fully connected layers has played a crucial role in preventing overfitting. Regarding model strategy, early stopping has significantly saved training time.

Furthermore, when examining the final results, it's evident that there is substantial room for improvement in the performance of the "earthworm" class, primarily due to the limited number of samples.

In summary, further tests are expected to fine-tune the model for better performance. Additionally, addressing the issue of poor performance in the "earthworm" class may involve using a more comprehensive and balanced dataset or refining the model itself.

6. Conclusion

In the field of agriculture, pest infestations often have adverse effects on the quality and quantity of crops. Therefore, this study compared four CNN models based on transfer learning, with a primary focus on utilizing pre-trained models to address the issue of pest classification. This study employed a transfer learning approach, using pre-trained models to tackle the pest problem. Additionally, an in-depth analysis of the impact of sample size and hyperparameters is conducted.

Experimental results have shown that the EfficientNet and Xception models exhibit excellent performance in pest classification, but the Xception model has a larger model size and training time. Overall, the EfficientNet model performed the best. In the pest dataset, the EfficientNet model achieved an accuracy of 88.26% in classification.

In summary, lightweight models perform better for pest classification, making them a more suitable choice for this relatively uncomplicated problem. Complex models do not show a significant accuracy advantage when dealing with simpler issues and also consume more computational resources and training time.

The future goal is to further refine the model, with particular emphasis on improving the performance of smaller pest categories, to enhance computational efficiency. Furthermore, this research has significant practical implications for assisting farmers in early pest identification and classification in agricultural fields.

References

- [1] Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017, arXiv preprint arXiv: 1704.04861.
- [2] Tan, Mingxing, and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning, 2019: 6105-6114.
- [3] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016: 770-778.
- [4] Chollet, François. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2017: 1251-1258.
- [5] Faithpraise, Fina, Philip Birch, Rupert Young, J. Obu, Bassey Faithpraise, and Chris Chatwin. Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters. *Int. J. Adv. Biotechnol*, 2013, 4(2): 189-199.
- [6] Huang, Mei-Ling, Tzu-Chin Chuang, and Yu-Chieh Liao. Application of transfer learning and image augmentation technology for tomato pest identification. *Sustainable Computing: Informatics and Systems*, 2022, 33: 100646.
- [7] T. Kasinathan, D. Singaraju, and S. R. Uyyala. Insect classification and detection in field crops using modern machine learning techniques. *Information Processing in Agriculture*, 2021, 8(3): 446-457.
- [8] Lu, Yang, Shujuan Yi, Nianyin Zeng, Yurong Liu, and Yong Zhang. Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, 2017, 267: 378-384.
- [9] Li, Chen, Tong Zhen, and Zhihui Li. Image classification of pests with residual neural network based on transfer learning. *Applied Sciences*, 2022, 12(9): 4356.
- [10] Ngugi, Lawrence C., Moataz Abelwahab, and Mohammed Abo-Zahhad. Recent advances in image processing techniques for automated leaf pest and disease recognition—A review. *Information processing in agriculture*, 2021, 8(1): 27-51.
- [11] Wang, RuJing, Jie Zhang, Wei Dong, Jian Yu, ChengJun Xie, Rui Li, TianJiao Chen, and HongBo Chen. A crop pests image classification algorithm based on deep convolutional neural network. *Telecommunication Computing Electronics and Control*, 2017, 15(3): 1239-1246.
- [12] Agricultural Pests Image Dataset. URL: <https://www.kaggle.com/datasets/vencerlanz09/agricultural-pests-image-dataset>. Last Access Time: 2023/11/11