

Article

An Advancing GCT-Inception-ResNet-V3 Model for Arboreal Pest Identification

Cheng Li ¹ , Yunxiang Tian ¹, Xiaolin Tian ^{1,*}, Yikui Zhai ^{2,*}, Hanwen Cui ^{3,4} and Mengjie Song ⁵

¹ Faculty of Innovation Engineering, School of Computer Science and Engineering, Macau University of Science and Technology, Macau 999078, China; 2009853vi011006@student.must.edu.mo (C.L.); 2109853gii30009@student.must.edu.mo (Y.T.)

² The Department of Intelligent Manufacturing, Wuyi University, Jiangmen 529020, China

³ State Key Laboratory of Lunar and Planetary Sciences, Macau University of Science and Technology, Macau 999078, China; hwcui@must.edu.mo

⁴ School of Computer Science, Zhuhai College of Science and Technology, Zhuhai 519041, China

⁵ College of Plant Protection, China Agricultural University, Beijing 100107, China; 2022319010410@cau.edu.cn

* Correspondence: xltian@must.edu.mo (X.T.); yikuizhai@163.com (Y.Z.)

Abstract: The significance of environmental considerations has been highlighted by the substantial impact of plant pests on ecosystems. Addressing the urgent demand for sophisticated pest management solutions in arboreal environments, this study leverages advanced deep learning technologies to accurately detect and classify common tree pests, such as “mole cricket”, “aphids”, and “*Therioaphis maculata* (Buckton)”. Through comparative analysis with the baseline model ResNet-18 model, this research not only enhances the SE-RegNetY and SE-RegNet models but also introduces innovative frameworks, including GCT-Inception-ResNet-V3, SE-Inception-ResNet-V3, and SE-Inception-RegNetY-V3 models. Notably, the GCT-Inception-ResNet-V3 model demonstrates exceptional performance, achieving a remarkable average overall accuracy of 94.59%, average kappa coefficient of 91.90%, average mAcc of 94.60%, and average mIoU of 89.80%. These results signify substantial progress over conventional methods, outperforming the baseline model’s results by margins of 9.1%, nearly 13.7%, 9.1%, and almost 15% in overall accuracy, kappa coefficient, mAcc, and mIoU, respectively. This study signifies a considerable step forward in blending sustainable agricultural practices with environmental conservation, setting new benchmarks in agricultural pest management. By enhancing the accuracy of pest identification and classification in agriculture, it lays the groundwork for more sustainable and eco-friendly pest control approaches, offering valuable contributions to the future of agricultural protection.

Keywords: arboreal pest identification; agriculture; deep learning; GCT-Inception-ResNet-V3 model; image processing



Citation: Li, C.; Tian, Y.; Tian, X.; Zhai, Y.; Cui, H.; Song, M. An Advancing GCT-Inception-ResNet-V3 Model for Arboreal Pest Identification. *Agronomy* **2024**, *14*, 864. <https://doi.org/10.3390/agronomy14040864>

Academic Editor: Francis Drummond

Received: 15 March 2024

Revised: 18 April 2024

Accepted: 18 April 2024

Published: 20 April 2024

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The criticality of effective pest management in safeguarding agricultural productivity and environmental health has heightened in recent years [1]. Consequently, understanding and mitigating tree pest threats has become an imperative endeavor. Tree pests, such as mole crickets, aphids, and *Therioaphis maculata* (Buckton) [2], detrimentally affect crop yields and forest ecosystems, underscoring the urgent need to devise sophisticated detection and management tactics. Trees, meanwhile, are pivotal to both the natural environment and human societies; they are foundational to ecosystems and crucial in sustaining biodiversity [3], climate regulation, soil and water conservation, and providing ecological services. Thus, the application of advanced technology to address arboricultural pest issues is imperative [4].

This study focuses on three prevalent arboricultural pests—“mole cricket”, “aphids”, and “*Therioaphis maculata* (Buckton)” —which pose significant risks to tree health and

ecosystem stability. Mole crickets are infamous for their burrowing activities that harm roots and compromise trees' structural stability. Conversely, aphids' sap consumption leads to diminished tree vitality, stunted growth, and increased disease vulnerability [5]. *Therioaphis maculata* (Buckton) inflicts severe leaf damage, affecting photosynthesis and overall tree health. This research employed cutting-edge deep learning technology for the automated classification and recognition of these arboricultural pests. Accurate pest identification is crucial for effective management and environmental preservation [6].

In the realm of agricultural pest detection and classification, researchers have leveraged a variety of datasets to train and test machine learning models in the previous studies, such as the Rice Pest Image Dataset, the PlantVillage Dataset and the Agricultural Pest Dataset (APD) [7–9]. Despite advancements in previous plant pest identification research, existing limitations persist. Conventional approaches largely depend on expert human judgment and manual observation, which are susceptible to inaccuracies and bias, resulting in flawed pest identification [10,11]. Furthermore, these traditional methods are not scalable and cannot keep up with the vast diversity and adaptability of pests across different regions and crops, making comprehensive monitoring and management a challenging task [12,13]. Automated systems using machine learning offer a promising solution, yet they require vast amounts of data and computational resources, often lacking in field conditions [14]. The ResNet architecture, with its deep residual learning framework [15,16], offers a potential improvement in learning complex features for accurate pest identification, yet like other advanced CNN (Convolutional Neural Network) architectures (e.g., VGG16 [17,18], DenseNet [19], and Inception-V3 [20–22]), it faces challenges in deployment on resource-constrained mobile devices, thus restricting their accessibility. Additionally, memory-efficient CNN architectures [23,24], although apt for mobile environments, typically sacrifice classification precision. The need for real-time processing and analysis in the field further exacerbates these limitations, highlighting the necessity for optimized models that balance performance and efficiency. These challenges necessitate a novel methodology that marries deep learning accuracy with pragmatic deployment solutions.

To address the aforementioned issues, this paper introduces three novel models: GCT-Inception-ResNet-V3, SE-Inception-ResNet-V3, and SE-Inception-RegNetY-V3 [25,26]. These models were compared against the traditional ResNet-18 model, which was selected as the control group. The selection of ResNet-18 as the baseline model is due to its widespread adoption in previous research on pest and disease identification [15–17], alongside SE-RegNet, SE-SE-RegNet, SE-RegNetY, and GCT-Inception-ResNet-V3. The GCT-Inception-ResNet-V3 model significantly outperforms the baseline model group. It achieves an overall accuracy of 94.59%, which is 9.1% higher. Its Kappa coefficient is 13.7% better, reaching 91.90%. The model's mean accuracy (mAcc) is 9.1% higher at 94.60%, and its mean Intersection over Union (mIoU) surpasses the control by 15%, achieving 89.80%. Essentially, this model shows marked improvements across all measured metrics, demonstrating higher precision and reliability.

The primary innovations of this study compared to prior research are:

1. The introduction of three innovative models, notably GCT-Inception-ResNet-V3, which outperform traditional deep learning models used in pest research [26]. The GCT-Inception-ResNet-V3 model, in particular, exhibits significant improvement over the baseline model. Also, the standard deviations of the indicators of the proposed model are all smaller than 0.001, demonstrating the stability and applicability of the model;
2. Previous tree-specific pest studies have been limited, often focusing on a single tree species, with a lack of deep learning-based pest research for a broader range of trees. This research targets a wide variety of trees, specifically arborvitae [27,28];
3. While prior research on tree insect pests primarily concentrated on intelligent target detection, there has been little emphasis on their identification and classification [29], which is the focal point of this study;

4. Incorporating the GCT (Gated Channel Transformation) attention mechanism, the GCT-Inception-ResNet-V3 model is demonstrated to be less time-consuming and more efficient compared to models utilizing the traditional SE (Squeeze-and-Excitation) Attention mechanism [26];
5. The model proposed in this study combines the improved Feature Extraction Algorithm to enhance the model so that it is also applicable to the identification of different pests in other different types of crops in agriculture, while the preprocessing part of the data contains image processing operations such as median filtering, histogram equalization, etc., which can greatly help in avoiding the detection of the effects of weather, sunlight, and soot, which are the most common weather conditions in agricultural fields. Therefore, this study will be applied to a wider range of agricultural crops and pests in the future.

The rest of this paper is organized as follows. Section 2 introduces the data and the models used in this study. Section 3 shows the experimental results. Section 4 offers discussions of the effectiveness of the proposed GCT-Inception-V3 model. Section 5 gives the conclusions of the paper.

2. Materials and Methodology

2.1. Dataset

The IP102 dataset, a specialized image collection for agricultural pests, is designed to foster the advancement of image recognition technology within this field [30]. Comprising images of 102 prevalent agricultural pests across various categories, such as insects and mites, this dataset underscores the substantial impact these pests have on agricultural productivity, posing significant threats to crop growth and yield. The IP102 dataset serves as a valuable resource for researchers, facilitating the application of machine learning and deep learning methodologies to devise an effective pest recognition and classification system. Such advancements aim to bolster agricultural production through intelligent pest management.

For this research, three arboreal pests from the IP102 dataset—mole cricket, aphids, and *Therioaphis maculata* (Buckton)—were chosen for investigation due to their pronounced relevance to environmental health. These pests are notorious in agricultural contexts, directly endangering crop vitality and yield. The dataset encompasses 1905 images of these pests, forming the training set (635 training samples for each of the three classes), and an additional 333 images constituting the test set to assess the model's pest recognition efficacy (the number of test sets for the mole cricket, aphid, and Buckton comprise 119, 107 and 107, respectively). Figure 1 illustrates samples from the pest dataset.

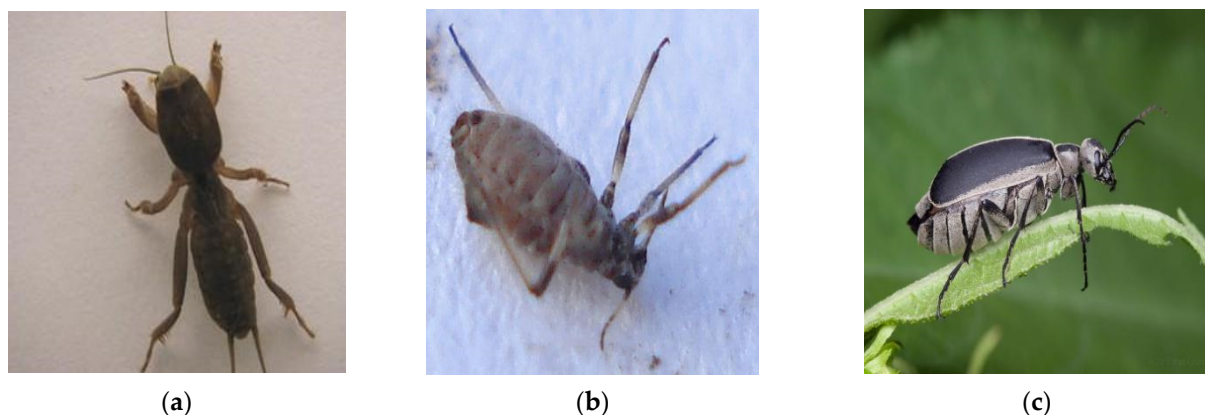


Figure 1. Pest dataset samples: (a) mole cricket; (b) aphids; (c) *Therioaphis maculata* (Buckton).

To augment the dataset and enhance the model's generalization capabilities, this study employs various data preprocessing techniques, including image flipping and rotation. These methods create new training samples from the original images, thereby increasing

data diversity [31]. Furthermore, the pixel dimensions of the images were standardized, with all images uniformly resized to 60×60 pixels. This adjustment not only diminishes the computational demands of model training, but also accelerates the training process while ensuring the model's effectiveness with input images of varying sizes [11]. Also in this study, histogram equalization has been performed to improve the overall contrast of the image, to make the details in the image more visible and to reduce the effect of strong light, weak light, brightness etc.; in this study, a median filter is used to reduce the noise in order to reduce the effect caused by dust, shadows, smoke etc., so as to efficiently remove the random noise in the image while maintaining the edges and details of the image [12]. Through these data enhancement and preprocessing measures, the research aims to develop an efficient and robust pest identification model that facilitates the rapid and precise detection of agricultural pests, offering technical support for pest management strategies. The processed RGB color image data are fed into the model for training and testing. Figure 2 illustrates the research's technical pathway.

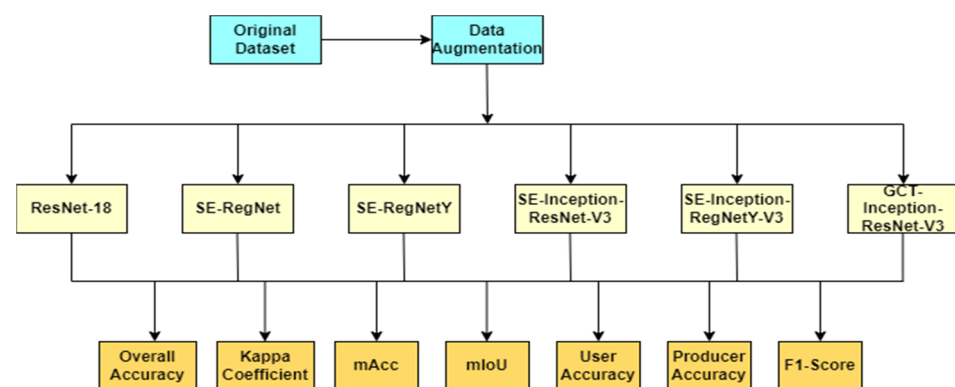


Figure 2. Technical route.

2.2. Methodology

2.2.1. ResNet-18 Model

The ResNet-18 model, a derivative of the Residual Network architecture, introduces a pioneering method for deep neural network training by incorporating skip connections, which circumvent certain layers [32]. These connections counteract the vanishing gradient issue, facilitating the training of significantly deeper networks than were previously achievable [33]. The ResNet-18 model, with its 18-layer deep structure, has demonstrated effectiveness in various image recognition tasks, attributed to its capacity to extract highly distinctive features from a broad spectrum of images while maintaining computational efficiency [34]. This makes ResNet-18 a suitable choice for the task of arboreal pest identification. Particularly, when identifying pests like mole cricket, aphids, and *Therioaphis maculata* (Buckton) from the IP102 dataset, the ResNet-18 model's ability to discern detailed patterns and textures is crucial [35].

For this study, the ResNet-18 model has been customized for classifying three common arboreal pests associated with environmental impact. The model starts with a convolutional layer of 3×3 kernel size, a stride of 1, and padding of 1, followed by batch normalization and ReLU activation, setting the stage for feature extraction. This layer is tailored to accommodate input images resized to 60×60 pixels, optimizing the model for this specific dataset. The model's architecture includes sequences of basic blocks, each containing two 3×3 convolutional layers with batch normalization and ReLU activation, arranged in four sequences with the configuration [2, 2, 2, 2], aligning with the ResNet-18 standard. To tailor the model for pest detection using the IP102 dataset, several adjustments and optimizations were applied, including a grid search for optimal parameters. To mitigate overfitting and boost generalization, a dropout rate of 0.5 was implemented before the final fully connected layer. Figure 3 illustrates the ResNet-18 model's adapted architecture used in this research.

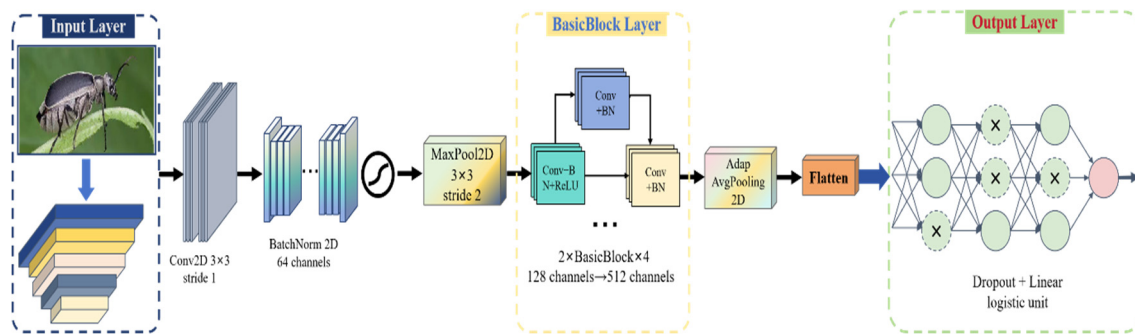


Figure 3. The architecture of the ResNet-18 Model.

2.2.2. SE-RegNet Model

The SE-RegNet model embodies a cutting-edge architectural integration, merging the strengths of Squeeze-and-Excitation (SE) blocks with the efficient RegNet framework, designed for precise and effective image classification tasks [36]. SE blocks, which adaptively recalibrate channel-wise feature responses, boost the network's representational capacity by methodically capturing interdependencies among channels, achieving notable performance enhancements with minimal additional computational demand [37]. RegNet, recognized for its straightforwardness and scalability, offers a modular structure that can be readily scaled and customized for various applications. The selection of SE-RegNet for arboreal pest identification is driven by the requirement for a model adept at processing the complex and nuanced imagery of pests, capable of accentuating pertinent features and diminishing irrelevant ones [38]. This function is vital for differentiating among similar pest species and variations within the IP102 dataset, where minor visual indicators are key for precise categorization. The integration of SE blocks into the RegNet architecture aims to exploit spatial and channel-wise attention mechanisms, thereby improving feature extraction and learning efficacy, positioning SE-RegNet as an optimal choice for this demanding field.

The SE-RegNet model is intricately designed to excel in classifying three common arboreal pests related to environmental concerns from the IP102 dataset. At the heart of its architecture is the incorporation of Squeeze-and-Excitation (SE) blocks, which introduce a dynamic channel-wise attention mechanism to enhance the feature learning process. These blocks employ global average pooling to condense spatial information into a succinct channel descriptor, which is subsequently refined through a series of fully connected layers, resulting in a channel-wise modulation of the feature maps based on the learned importance of each channel. Figure 4 illustrates the architectural configuration of the SE-RegNet model.

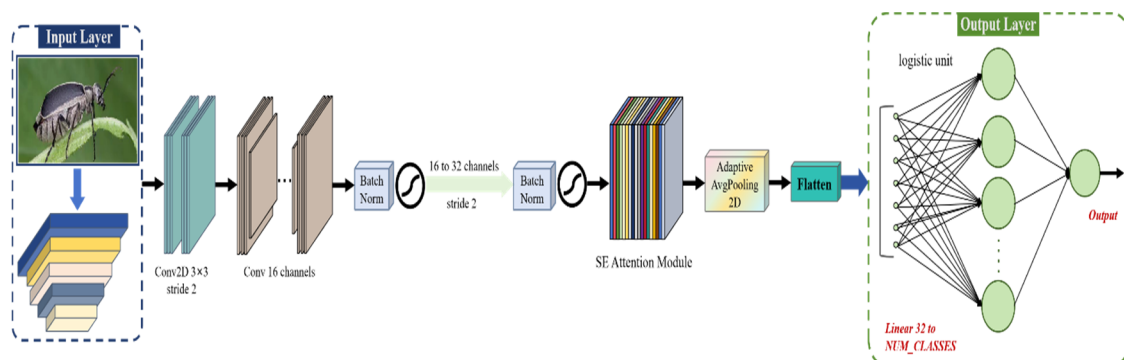


Figure 4. The architecture of the SE-RegNet Model.

2.2.3. SE-RegNetY Model

The SE-RegNetY model marks a significant advancement in deep learning for image classification, particularly tailored for the intricate demands of environmental monitoring, such as in arboreal pest detection [39]. This model stands out by integrating the dynamic,

channel-wise attention modulation of the Squeeze-and-Excitation (SE) mechanism with the structural efficiency of the RegNet architecture, forming the specialized SE-RegNetY framework [40]. The incorporation of the SE mechanism within the RegNet framework facilitates a nuanced, focused recalibration of features, ensuring the model accentuates the most relevant features for classification tasks. This feature is crucial in pest recognition, where minor distinctions between species are essential for precise identification [41]. The SE-RegNetY model, distinct from conventional SE-RegNet models, is specifically designed to offer enhanced precision and adaptability, addressing the complex variability found in natural settings, such as changing light conditions, various pest positions, and backgrounds that challenge typical classification models.

The SE-RegNetY model is intricately constructed using YBlocks as its fundamental components, each consisting of a convolutional layer followed by batch normalization and ReLU activation. This configuration constitutes the model's backbone, facilitating efficient feature extraction from input images resized to 60×60 pixels. The distinctive aspect of the YBlock is the incorporation of SE-Blocks, which dynamically modulate the network's channel weights, enabling the model to focus on features crucial for pest classification. This modulation is accomplished via adaptive average pooling within the SE-Block, which condenses spatial information into a channel descriptor, itself subsequently utilized to adjust the feature maps, thereby amplifying the network's representational capacity. The Adam optimizer is utilized, incorporating a weight decay rate for regularization, to refine the training process, ensuring consistent progress. Moreover, L2 regularization is implemented to mitigate overfitting. Figure 5 depicts the SE-RegNetY model's architecture.

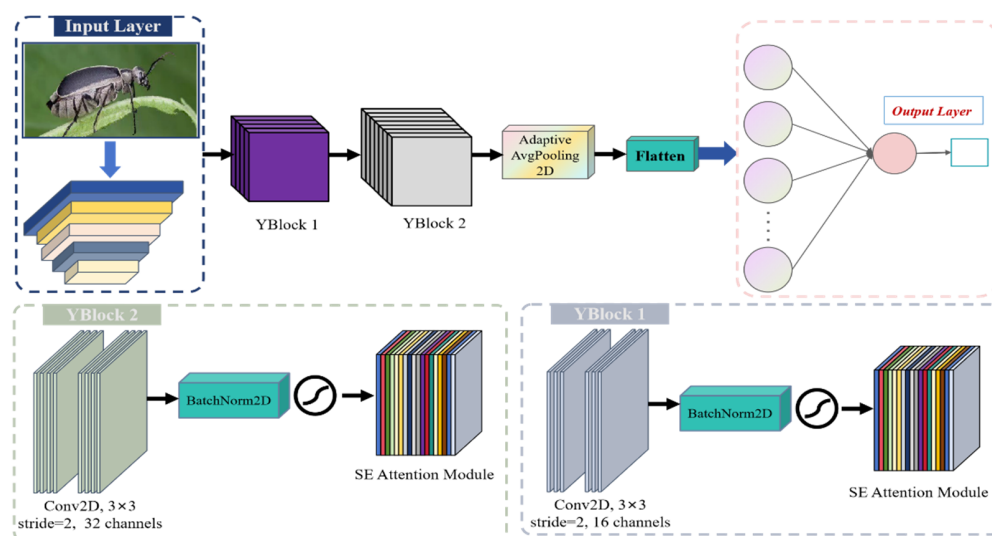


Figure 5. The architecture of the SE-RegNetY model.

2.2.4. SE-Inception-ResNet-V3 Model

The SE-Inception-ResNet-V3 model represents an advanced integration of three key architectural concepts in deep learning: Squeeze-and-Excitation (SE) blocks, Inception modules, and the Residual Network (ResNet) framework. This composite architecture is deliberately selected for arboreal pest detection, attributed to its superior capacity to identify and accentuate pivotal features within intricate visual data. The inclusion of SE blocks facilitates dynamic channel-wise recalibration, augmenting the model's acuity in recognizing essential characteristics that distinguish various pest species [42]. Inception modules enhance the model's adaptability in managing diverse spatial dimensions, allowing the effective analysis of detailed pest imagery in natural settings [43]. The design of this model is particularly apt for detecting specific arboreal pests from the IP102 dataset, where the visual resemblances among different species present a notable classification challenge. The differences in appearance, scale, and context of these pests in the images demand a model proficient in detecting subtle nuances while ensuring high accuracy levels.

The SE-Inception-ResNet-V3 model is meticulously engineered to optimize its performance in analyzing and classifying complex image data. The network's initial phase comprises a convolutional layer with a 7×7 kernel, a stride of 2, and padding of 3, followed by batch normalization and ReLU activation. This configuration effectively reduces the input's spatial dimensions while initiating the feature extraction process. A subsequent max-pooling layer further downsamples the input, setting the stage for the ensuing layers. At the core of the model are the modified Basic Blocks that integrate SE functionality, enabling the network to dynamically recalibrate its focus on the most salient features. After the initial residual blocks, the model progresses to a series of Inception modules, which divide the input into various branches processed with distinct kernel sizes (1×1 , 5×5 , and 3×3) before concatenating the results. This branched strategy allows the model to capture an extensive spectrum of spatial information, encompassing both fine details and larger contextual elements. The incorporation of L2 regularization and a dropout layer further bolsters the model's robustness and generalizability. The distinctive architecture of the model is illustrated in Figure 6.

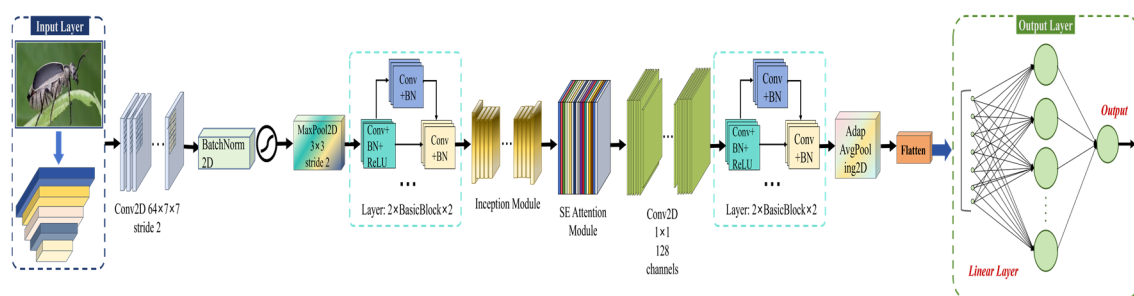


Figure 6. The architecture of the SE-Inception-ResNet-V3 model.

2.2.5. SE-Inception-RegNetY-V3 Model

The SE-Inception-RegNetY-V3 model represents a state-of-the-art architecture tailored for the intricate task of arboreal pest classification, harnessing the advantages of Squeeze-and-Excitation (SE) blocks, Inception modules, and the RegNet framework. This amalgamation responds to the necessity for a model capable of adeptly managing the significant variability and complexity found in environmental imagery, particularly within the IP102 dataset for pest identification. SE blocks introduce an adaptive mechanism for recalibrating channel-wise feature responses, markedly improving the model's capacity to emphasize pertinent features [44]. Inception modules facilitate the extraction of information across multiple scales, enabling the model to discern detailed attributes of pests and their environments—a critical aspect for differentiating similar species [45]. Additionally, the RegNet backbone provides a scalable and efficient structure, supporting deep learning processes while keeping computational demands in check. The selection of the SE-Inception-RegNetY-V3 model for this task is based on its superior proficiency in addressing the nuanced differences and high intraclass variation among pest species. The model's dynamic attention mechanism, courtesy of SE blocks, along with the multi-scale feature analysis afforded by Inception modules, guarantees thorough feature extraction [46].

The SE-Inception-RegNetY-V3 model's architecture is meticulously designed to optimize feature extraction and classification efficacy. The network begins with a convolutional layer that readies the input for advanced processing, followed by the implementation of Basic Blocks enhanced with SE functionality. This early incorporation of attention mechanisms primes subsequent layers to focus on pertinent features, reducing the influence of background noise. At its core, the model boasts a series of Inception-YBlocks, a novel amalgamation that combines YBlock architecture with Simplified Inception Modules. These blocks are structured to analyze inputs via parallel pathways, capturing both local and broad contexts through the use of various kernel sizes. Embedding SE-Blocks within these Inception-YBlocks further hones the feature maps, accentuating vital details and minimizing redundancies. The model adopts a weight decay approach during optimization to

enhance learning regularization, promoting adaptability to novel data. Figure 7 displays the SE-Inception-RegNetY-V3 model's architecture.

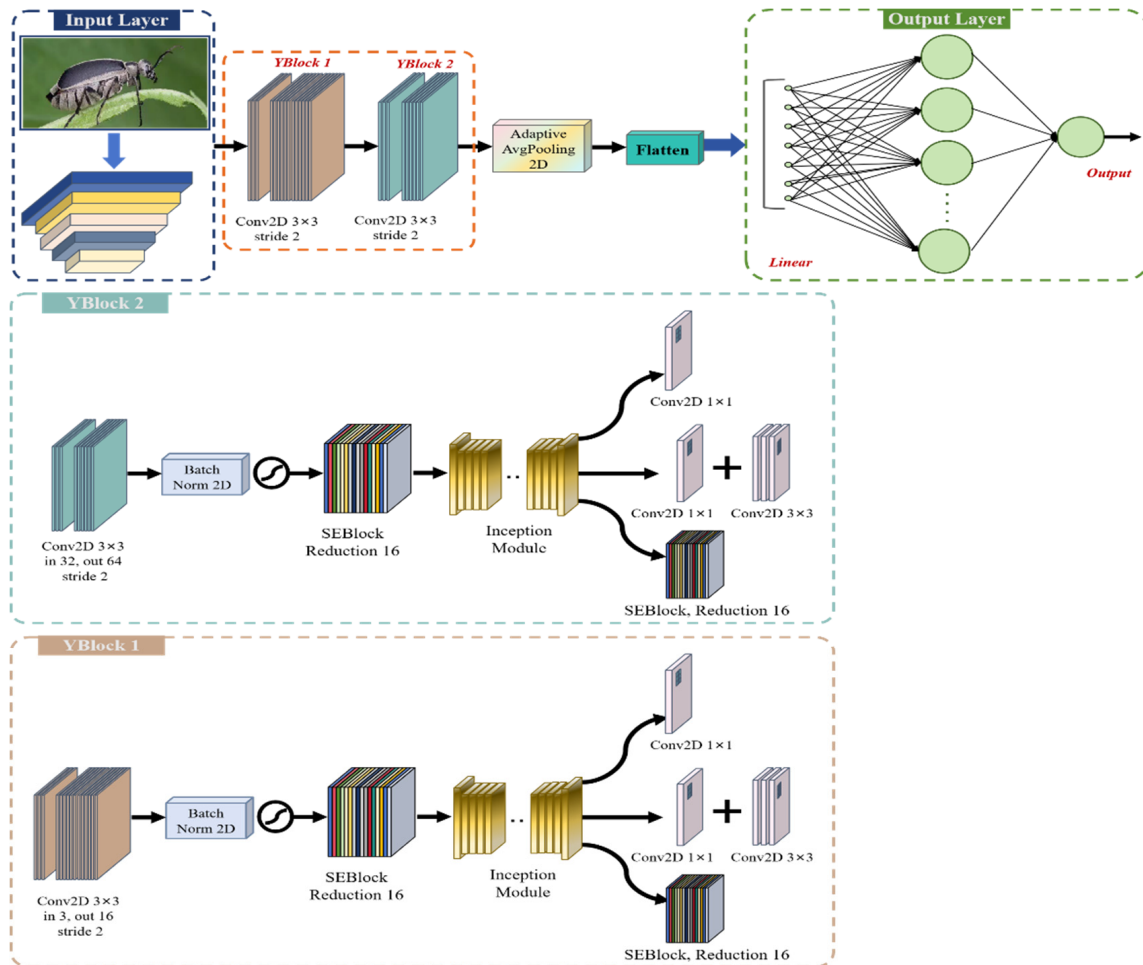


Figure 7. The architecture of the SE-Inception-RegNetY-V3 model.

2.2.6. GCT-Inception-ResNet-V3 Model

The GCT-Inception-ResNet-V3 model represents a groundbreaking architecture specifically designed for the complex task of arboreal pest detection. It skillfully merges Global Context Transformation (GCT) mechanisms with the adaptability of Inception modules and the robustness of the ResNet framework, and thus creates a formidable solution for deep learning-driven image classification [26]. The GCT mechanism stands out for its use of channel-wise feature normalization to boost the network's representational efficiency. By adjusting features according to their global statistical characteristics, GCT ensures the model accentuates pertinent patterns while minimizing noise, thus enhancing focus and discriminability. This model is chosen for arboreal pest identification due to several key reasons. First, the GCT's capacity to refine feature representation is ideally suited to the subtle distinctions among various pest species, where minor details can greatly influence classification outcomes. Second, the multi-scale processing ability of the Inception modules enables the network to discern a broad array of spatial features, ranging from fine details to larger shapes, vital for detecting pests of varying sizes and orientations. Finally, integrating these elements within a ResNet backbone facilitates deep feature extraction while preventing vanishing gradients, courtesy of the residual connections that promote effective learning in deep networks. Consequently, the GCT-Inception-ResNet-V3 model is an exemplary choice for the complex requirements of pest recognition within natural environments. The sophisticated management of global and local features, paired with

the dynamic adjustment of feature channels, establishes a new standard for accuracy and robustness in environmental image classification tasks. Within convolutional networks, let x belong to $\mathbb{R}^{C \times H \times W}$, denoting an activation feature with H and W as spatial dimensions and C as the channel count. GCT operates via the following transformation, shown in Formula (1):

$$\hat{x} = F(x|a, \gamma, \beta), \quad a, \gamma, \beta \in \mathbb{R}^C, \quad (1)$$

where the embedding weight a adapts outputs, and the gating weight γ along with bias β control gate activation, influencing GCT's channel-specific actions. Importantly, GCT's parameter load, $O(C)$, demonstrates greater efficiency compared to SE Attention module's $O(C^2)$ [26].

The GCT-Inception-ResNet-V3 model's architecture is intricately crafted to optimize both efficiency and effectiveness in processing high-dimensional image data. Initiating with a BasicConv2d layer featuring a 7×7 convolution, the model kickstarts the feature extraction process, which is then refined by max-pooling to diminish spatial dimensions while retaining vital information. At the heart of the architecture lies the integration of Inception modules, which divide the input into multiple branches, each processed with different kernel sizes, enabling the network to concurrently capture information across various scales. This feature is essential for processing complex images where pests are set against varied backgrounds and orientations. Subsequent to the Inception module, a GCT layer globally normalizes features, boosting the model's emphasis on significant patterns via a transformation that involves scaling and shifting parameters. Additional convolutional layers and an adaptive average pooling layer follow, further honing the features before classification. A fully connected layer then conducts the final classification, mapping the processed features to specific pest categories. The GCT-Inception-ResNet-V3 model represents an exemplary fusion of advanced neural network technologies, enhancing its proficiency in distinguishing closely related pest species. Through the strategic amalgamation of GCT for feature modulation, Inception modules for multi-scale processing, and ResNet for deep learning efficiency, the model establishes a new benchmark in environmental image analysis, marking notable progress in agricultural and environmental monitoring. The incorporation of a weight decay rate in the optimizer aids in regularizing the learning process, fostering a robust model capable of high precision across varied environmental settings. Figure 8 illustrates the GCT-Inception-ResNet-V3 model's architecture.

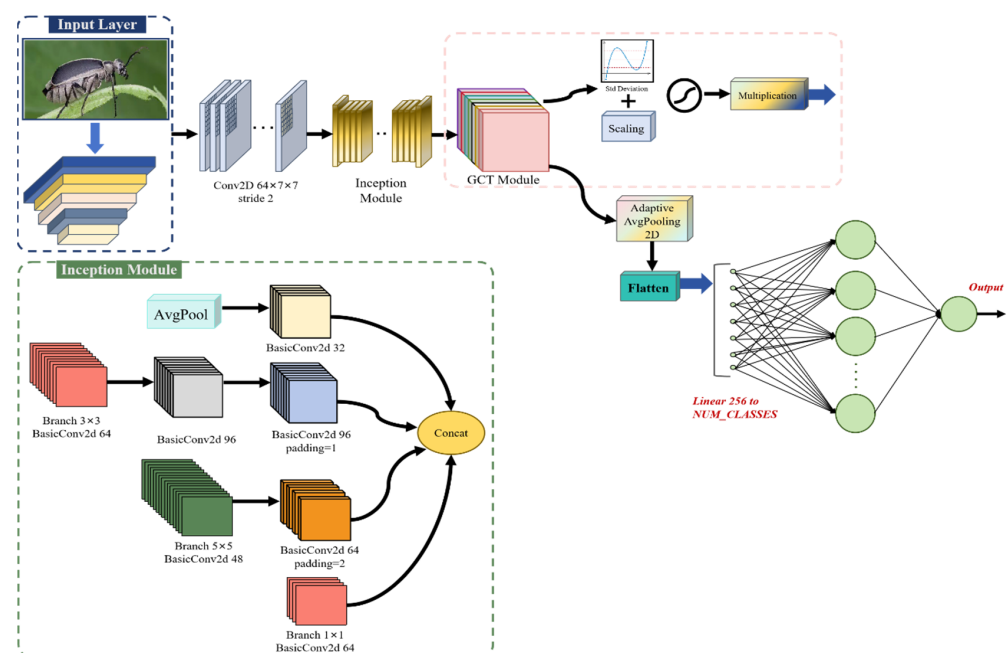


Figure 8. The architecture of the SE-Inception-RegNetY-V3 model.

2.3. Improved Algorithm Combination

The Enhanced Feature Extraction Algorithm, denoted as Algorithm 1, presents an intricate methodology for characterizing arboreal images to improve the identification of arboreal pests. This advanced approach includes morphological operations such as opening and top-hat transformations, adeptly engineered to highlight minute textural features and contrasts within the foliage, which are indicative of pest activity [47].

Algorithm 1 Improved Feature Extraction Algorithm

```

1: for X = 1 to 100 do
2: Image <- Read arboreal pest image
3: SegmentedRegion <- Segment the region to use in Image
4: DilatedRegion <- Dilate SegmentedRegion
5: ClosedRegion <- Apply closing operation to DilatedRegion
6: OpenedRegion <- Apply opening operation to ClosedRegion
7: TopHatTransformedRegion <- Apply top-hat transformation to OpenedRegion
8: ErodedRegion <- Apply erosion to TopHatTransformedRegion
9: EdgeImage <- Show image with eroded edges
10: end for

```

By amalgamating this algorithm with six distinct models, each specifically crafted for the detection of tree pests, the collective leverage optimizes the unique strengths inherent within diverse analytical frameworks [48]. The incorporation of the algorithm for feature extraction due to its robustness against variations in different situation and computational efficiency, crucial for the real-time detection of pests.

The Enhanced Feature Extraction Algorithm boasts several advantages:

1. Increased sensitivity—The preliminary application of morphological operations amplifies subtle textural discrepancies, potentially flagging the early stages of pest invasion, thus enabling timelier and more precise detection;
2. Improved feature relevance—The Dilation operation increases the size of the pest region, thus compensating for some of the edge detail that may be lost during segmentation [49];
3. Enhanced robustness—The top-hat transformation is specially designed to accentuate novel elements within the image, such as the manifestation of pests against complex backgrounds, often overlooked by conventional methodologies;
4. Flexibility—The algorithm is model-agnostic, thus facilitating seamless incorporation across various computational models to enhance their discernment of pest-afflicted regions within tree canopies;
5. Efficiency—Despite its complexity, the algorithm remains computationally economical, making it suitable for deployment in resource-limited environments, such as mobile applications or infield diagnostic tools.

The integration of Algorithm 1 into each of the six models is predicted to significantly elevate the precision of arboreal pest detection. This enhancement will yield models that are not merely more attuned to the different expressions of pest damage, but are also equipped to function consistently across diverse environmental settings, assuring robust and dependable pest detection.

3. Experimental Results

3.1. Experimental Setup

The experimental component of this study was conducted using Python version 3.7 and the PyTorch framework. The experiments were carried out on hardware equipped with NVIDIA RTX 3090 graphics cards. To ensure the reliability and stability of our findings, each model underwent ten separate experimental runs for comparison purposes. This approach allowed for a more comprehensive assessment of the models' performance and stability across different experimental conditions.

3.2. Setting of Hyperparameters

In this research, we adopted the ResNet-18 model as the baseline for pest identification studies due to its prominence in previous research [15,16]. The hyperparameters for the ResNet-18 were meticulously adjusted to suit the IP102 dataset's requirements, implementing a grid search for the optimal set. The network was configured with a batch size of 30 and a learning rate of 0.007, and it utilized the Adam optimizer to efficiently handle sparse gradients in complex environments. To enhance model generalization and prevent overfitting, a dropout rate of 0.5 was applied prior to the final fully connected layer.

For the SE-RegNet model, a grid search was employed to determine the best hyperparameters, settling on a learning rate of 0.0075, a batch size of 27, and an extended training duration of 270 epochs to achieve consistent model convergence. The Adam optimizer was selected for its dynamic adjustment capabilities of learning rates, complemented by dropout techniques to navigate the model through varied pest classification challenges effectively. The SE-RegNetY model also utilized a grid search for hyperparameter optimization, opting for a learning rate of 0.0075, a batch size of 23, and a training span of 260 epochs to ensure a thorough learning process while maintaining rapid convergence. It incorporated a weight decay rate and L2 regularization to further refine the training procedure and control overfitting. For the SE-Inception-ResNet-V3 model, the hyperparameters were fine-tuned through a grid search, selecting a learning rate of 0.007, a batch size of 23, and a succinct training period of 75 epochs to balance efficiency and effectiveness. The addition of L2 regularization and a dropout layer was intended to enhance the model's durability and generalization capabilities. The SE-Inception-RegNetY-V3 model's hyperparameters, identified via grid search, included a learning rate of 0.008, a batch size of 21, and a 200-epoch training regimen, carefully chosen to prevent overfitting while ensuring swift model convergence. The optimization process was augmented with a weight decay strategy to boost the regularization of the learning curve. The GCT-Inception-ResNet-V3 model's hyperparameter configuration was determined through a grid search, adopting a learning rate of 0.008 and a batch size of 29. This setup aimed to balance quick convergence with the minimization of overfitting risks, employing a weight decay rate in the optimizer to regularize the learning process effectively. These strategic selections and modifications across models demonstrate a comprehensive approach to optimizing pest detection performance while addressing the challenges of model training and generalization.

3.3. Results Show

This study assesses the efficacy of various models, including the standard ResNet-18, employing metrics such as overall accuracy (OA), Kappa coefficient, mAcc, and mIoU, to understand their effectiveness in pest identification [38]. Specifically, OA quantifies the percentage of correct predictions, computed as the total of true positives and true negatives divided by all observations, as delineated in Formula (2):

$$OA = (TP + TN) / (TP + TN + FP + FN) \quad (2)$$

Here, TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. The Kappa coefficient (k) quantifies the concordance between model predictions and actual observations, correcting for random agreement, as illustrated in Formula (3):

$$\text{Kappa Coefficient} = (Po - Pe) / (1 - Pe) \quad (3)$$

Mean Accuracy (mAcc) represents the average classification accuracy across different classes, providing insight into the model's uniformity in performance, particularly valuable in instances of dataset imbalance. The mAcc formula is:

$$mAcc = 1/N \times \sum_{i=1}^N TP_i / (TP_i + FN_i) \quad (4)$$

Mean Intersection over Union (mIoU) is crucial for segmentation tasks and evaluating the congruence between predicted and actual classifications, reflecting the model's precision in defining class boundaries and specifics, which is vital for tasks demanding meticulous accuracy:

$$\text{mIoU} = 1/N \times \sum_{i=1}^N TP_i / (TP_i + FP_i + TN_i) \quad (5)$$

Po signifies observed agreement, whereas Pe denotes the probability of random agreement. For a more comprehensive analysis, User's Accuracy (UA, or Recall), Producer's Accuracy (PA, or Precision), and the F1-Score are employed to examine the precision of each category within the models. UA evaluates the likelihood that a pixel classified in the map/image accurately represents its real-world class, highlighting the model's precision in identifying specific classes—a key metric for tasks requiring accurate class identification [37]. UA is calculated as the ratio of correct predictions (TP) to all predictions for that class (TP and FP):

$$\text{UA} = \text{TP} / (\text{TP} + \text{FP}) \quad (6)$$

PA, or Precision, indicates the likelihood that a pixel correctly classified corresponds to its true class, which is pivotal for reducing false positives. It sheds light on a model's accuracy in making positive predictions for each class, determined by dividing the true positives (TP) by the sum of true positives and false negatives (TP and FN) [50]:

$$\text{PA} = \text{TP} / (\text{TP} + \text{FN}) \quad (7)$$

The F-1 score integrates precision and recall into a singular metric that underscores their equilibrium. It is critical in scenarios where both precision and recall are equally important, ensuring one is not optimized at the expense of the other. This metric offers a comprehensive evaluation of model performance, as illustrated in Formula (8):

$$\text{F1-Score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (8)$$

These metrics elucidate the model's effectiveness, highlighting its strengths and areas for enhancement in pest identification.

Figure 9 illustrates the average value of the performance metrics—OA, Kappa coefficient, mAcc, and mIoU—of seven different models of the 10 separate experimental runs, including the baseline model ResNet-18. An in-depth analysis of these metrics for the evaluated models provides critical insights into their capabilities in classifying arboreal pests. The data indicate that the GCT-Inception-ResNet-V3, SE-Inception-ResNet-V3, SE-Inception-RegNetY-V3, and SE-RegNet models outperform the baseline model ResNet-18, achieving superior OA, Kappa coefficient, mAcc, and mIoU scores. Specifically, the average OA values for GCT-Inception-ResNet-V3, SE-Inception-ResNet-V3, SE-Inception-RegNetY-V3, SE-RegNet, and ResNet-18 are 0.9459, 0.8851, 0.8682, 0.8576, and 0.8547, respectively. The average Kappa coefficients for these models are 0.9190, 0.8280, 0.8020, 0.7860, and 0.7820, respectively. The average mAcc values follow the same order at 0.9460, 0.8854, 0.8684, 0.8573, and 0.8549, while the average mIoU scores are 0.8980, 0.7930, 0.7680, 0.7465, and 0.7475, respectively. These enhanced metrics underscore the superior capacity of these models to detect and categorize the intricate features of arboreal pests, signifying a notable progression beyond the ResNet-18 model. Notably, the GCT-Inception ResNet-V3 model excels across all four average accuracy indicators, showcasing improvements of 9.1% in OA, 13.7% in the Kappa coefficient, and 9.1% in mAcc, with an mIoU increase of 15% compared to the baseline model. For the GCT-Inception-ResNet-V3 model, the standard deviations of OA, Kappa coefficient, mAcc, and mIoU for the ten sets of experiments are 0.0003, 0.0006, 0.0003, and 0.0009, respectively, which are all less than 0.001, showing the stabilizing and statistical significance of this model. This substantial progress empha-

sizes the GCT-Inception ResNet-V3 model's precision and consistency in identifying arboreal pests.

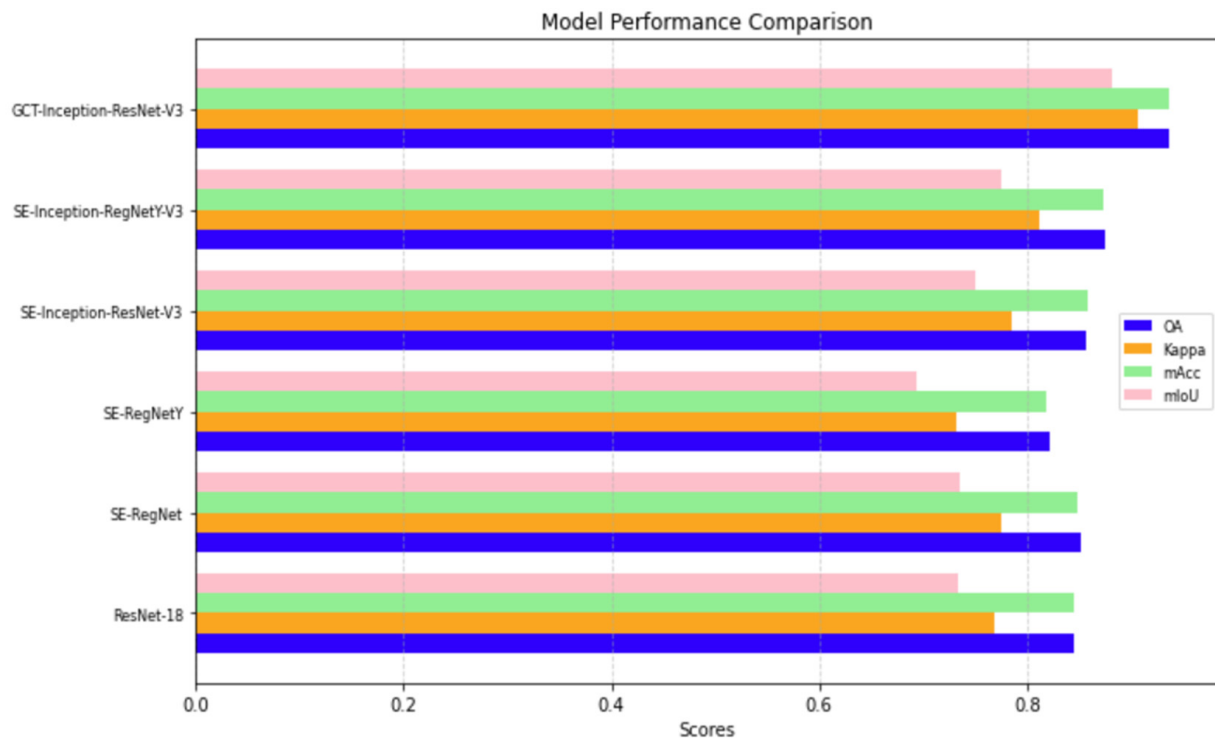


Figure 9. Model performance comparison.

OA, Kappa Coefficient, mAcc, and mIoU provide a comprehensive overview of a model's predictive accuracy and its alignment with empirical observations. The GCT-Inception ResNet-V3 model's exemplary scores in these metrics underscore its exceptional predictive proficiency and classification accuracy, thereby demonstrating its advanced capabilities in pest identification [51]. Nevertheless, for a holistic understanding, it is imperative to also incorporate average User's and Producer's Accuracies [52], as well as the F1-Score, in order to thoroughly assess the model's performance on specific pest characteristics, as delineated in Table 1. GCT-Inception ResNet-V3 demonstrates outstanding results in almost all aspects of average producer and user accuracies, and the average F1-score, affirming its leading status. The average performance matrix for each model, which offers detailed comparisons, is depicted in Figure 10. Figure 11 illustrates the variation in loss and validation set accuracy across epochs during the training phase of each model, wherein the GCT-Inception ResNet-V3 model reduces the loss to 0.0952 by the final epoch, and achieves a validation set accuracy of 0.9659.

Table 1. The UA (Recall), PA (Precision) and F1-Score of each category.

	Mole Cricket UA	Mole Cricket PA	Mole Cricket F1-Score	Aphids UA	Aphids PA	Aphids F1-Score	Bucton UA	Bucton PA	Bucton F1-Score
ResNet-18	0.8283	0.9011	0.8632	0.8980	0.8627	0.8800	0.8384	0.8058	0.8218
SE-RegNet	0.9596	0.8190	0.8837	0.9286	0.8349	0.8792	0.6837	0.9571	0.7976
SE-RegNetY	0.8980	0.8462	0.8713	0.8889	0.8148	0.8502	0.7041	0.8313	0.7624
SE-Inception-ResNet-V3	0.9192	0.8750	0.8966	0.9490	0.8378	0.8900	0.7879	0.9630	0.8667
SE-Inception-RegNetY-V3	0.8283	0.9011	0.8632	0.9184	0.8824	0.9000	0.8586	0.8252	0.8416
GCT-Inception-ResNet-V3	0.9596	0.9794	0.9694	0.9293	0.9388	0.9340	0.9490	0.9208	0.9347

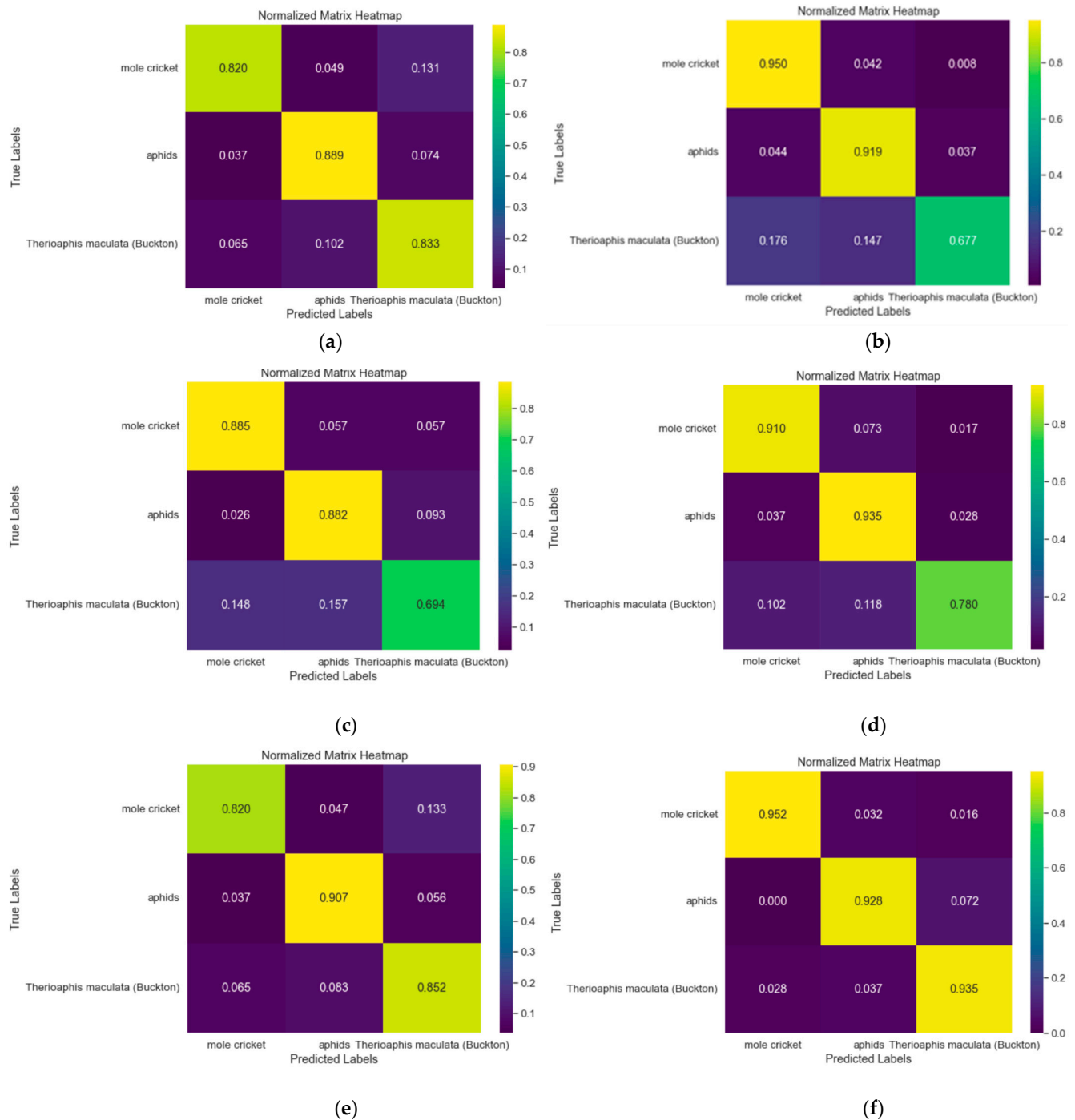
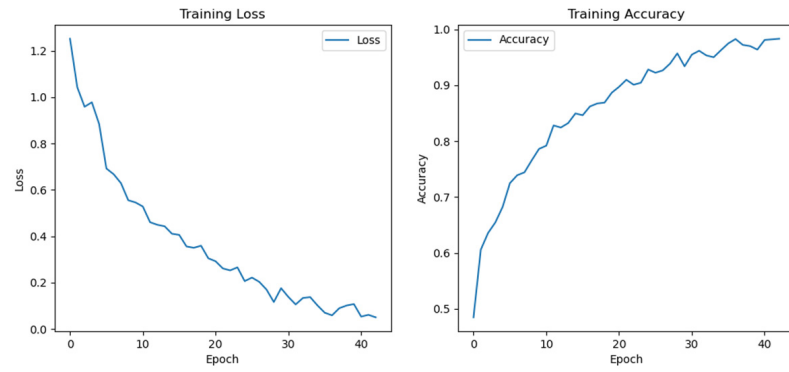
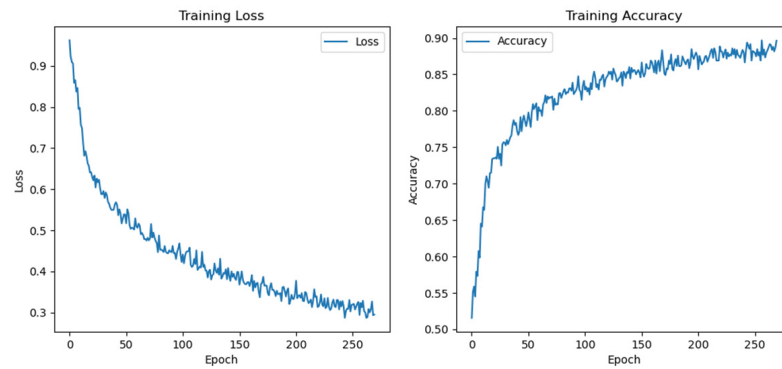


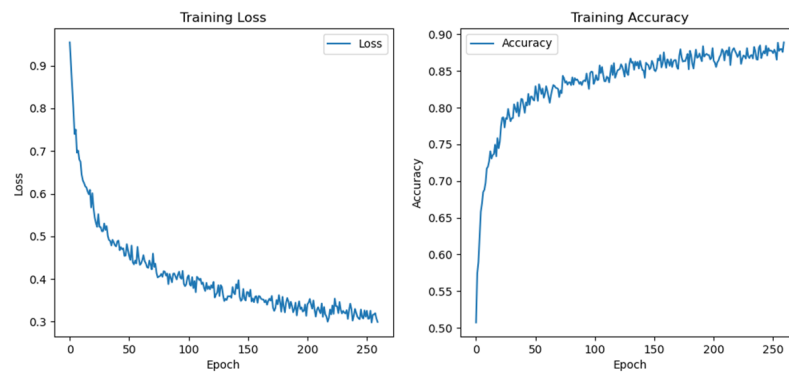
Figure 10. Normalized Matrix: (a) Normalized Matrix of ResNet-18; (b) Normalized Matrix of SE-RegNet; (c) Normalized Matrix of SE-RegNetY; (d) Normalized Matrix of SE-Inception-ResNet-V3; (e) Normalized Matrix of SE-Inception-RegNetY-V3; (f) Normalized Matrix of GCT-Inception-ResNet-V3.



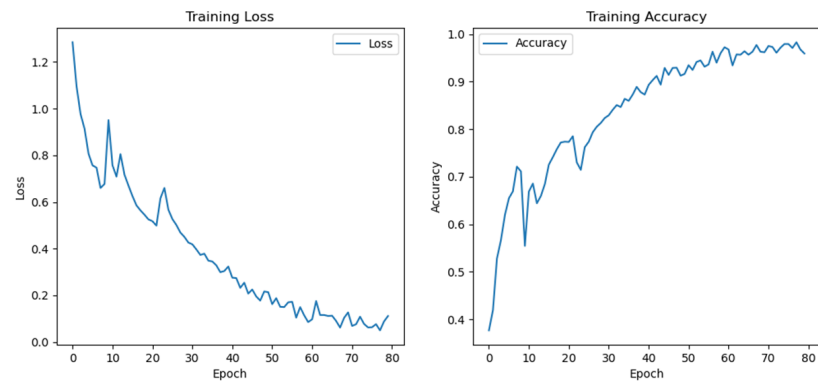
(a)



(b)



(c)



(d)

Figure 11. Cont.

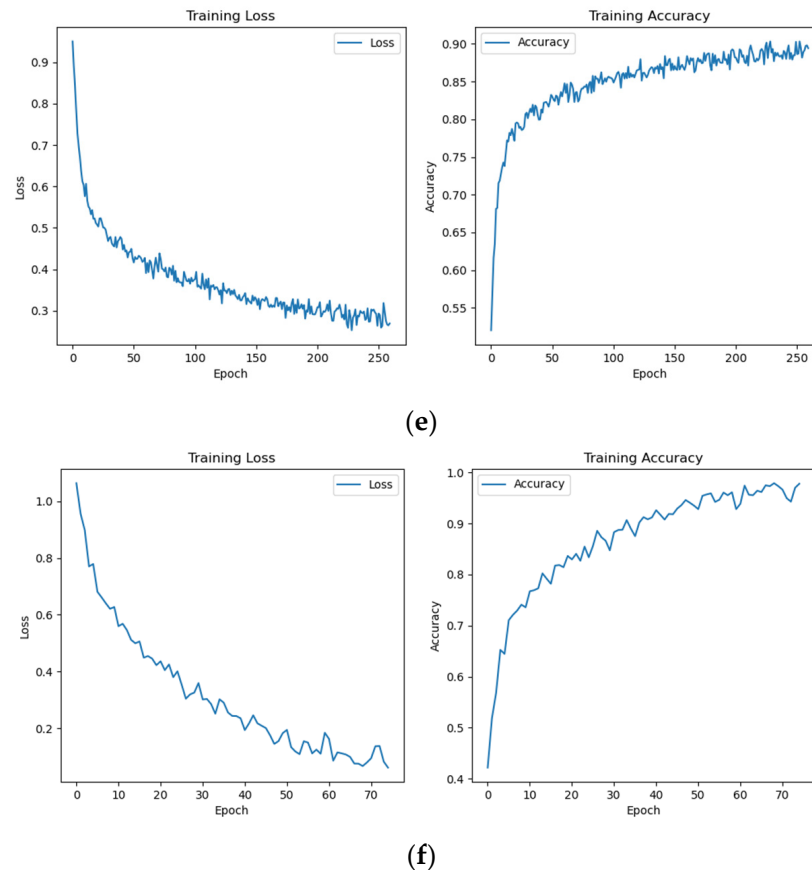


Figure 11. Loss and accuracy variation curve: (a) curve of ResNet-18; (b) curve of SE-RegNet; (c) curve of SE-RegNetY; (d) curve of SE-Inception-ResNet-V3; (e) curve of SE-Inception-RegNetY-V3; (f) curve of GCT-Inception-ResNet-V3.

4. Discussion

By analyzing the results above, the GCT-Inception-ResNet-V3 model achieved good results in terms of average OA, Kappa coefficient, mAcc, mIoU, as well as the UA and PA of each category in the ten groups of experiments, and the standard deviations of the evaluation indexes also reflect the stability and strong reliability of the model. To illustrate the convolutional neural network's (CNN) focus, a heat map is generated by capturing the gradients of the target class relative to the input image via backpropagation. A hook is registered at the final convolutional layer to collect the gradients, which are then applied to the layer's output to produce activation maps. These maps are subsequently weighted by the gradients' global average to generate a class activation map (CAM) [53]. The CAM is processed to emphasize areas critical to the model's decision-making, illustrating where the model concentrates its attention for classification. The CNN's visualization heat map is presented in Figure 12. This heat map is vital for discerning the features considered significant by the model, offering insights into its classification rationale. It is created by superimposing the CAM onto the original image, illuminating the model's focus areas, such as key body parts, characteristic points of pests, and the damage parts of the plants, which correspond with manual detection points and exhibit the GCT-Inception ResNet-V3 model's advanced learning capabilities.

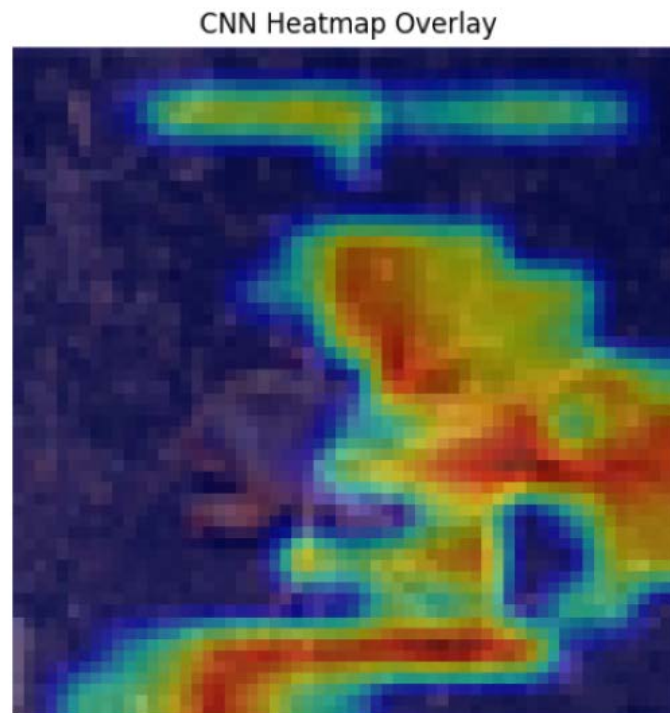


Figure 12. An exemplary result of the CNN visualization heat map.

5. Conclusions

This study was dedicated to the identification of arboreal pests, a critical aspect of supporting agricultural productivity and environmental conservation [54]. To this end, we explored and assessed the efficacy of various advanced models, including GCT-Inception-ResNet-V3, SE-Inception-ResNet-V3, and SE-Inception-RegNetY-V3, in addition to employing SE-RegNet and SE-RegNetY, with the well-established ResNet-18 model serving as the baseline for our comparative analysis. Among these, the GCT-Inception-ResNet-V3 model stood out, demonstrating exceptional performance by achieving an overall accuracy (OA) of 94.59%, a Kappa Coefficient of 91.90%, a mean accuracy (mAcc) of 94.60%, and a mean Intersection over Union (mIoU) of 89.80%. These results not only highlight the model's accuracy and reliability in identifying arboreal pests, but also validate the method's effectiveness. Additionally, the application of visualized convolutional neural network heatmaps provided insightful data on the model's focus areas, further affirming its practical utility and validity [53]. The study encountered certain limitations, including the models' varying degrees of sensitivity to different types of arboreal pests and the inherent challenges in generalizing the findings across diverse agricultural contexts. Future research directions will concentrate on developing innovative modeling approaches that can accommodate the complexities inherent in the investigation of arboreal pests and diseases. The aim is to refine and enhance agricultural conservation strategies further, contributing to the sustainability of agricultural practices and the preservation of environmental health [55].

Author Contributions: Conceptualization, C.L.; methodology, C.L.; software, C.L., Y.T. and M.S.; validation, C.L.; formal analysis, C.L. and Y.T.; investigation, C.L.; resources, C.L. and M.S.; data curation, C.L. and M.S.; writing—original draft preparation, C.L. and Y.T.; writing—review and editing, C.L. and X.T.; visualization, C.L. and H.C.; supervision, X.T., Y.Z. and H.C.; project administration, X.T. and Y.Z.; funding acquisition, X.T. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Wuyi University Hong Kong Macao Joint Research and Development Fund under Grant 2022WGALH19, and in part by the Science and Technology Development Fund of Macau (grant number 0038/2020/A1).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data of this study were obtained from the publicly available dataset IP102, accessed at <https://github.com/xpwu95/IP102> Accessed on 13 January 2024.

Acknowledgments: This work was supported by Wuyi University and Macau University of Science and Technology.

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

References

1. Loti, N.N.A.; Noor, M.R.M.; Chang, S.W. Integrated analysis of machine learning and deep learning in chili pest and disease identification. *J. Sci. Food Agric.* **2021**, *101*, 3582–3594. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Guo, Y.; Gao, J.J.; Wang, X.F.; Jia, H.Y.; Wang, Y.A.; Zeng, Y.; Tian, X.; Mu, X.Y.; Chen, Y.; Ouyang, X. Precious Tree Pest Identification with Improved Instance Segmentation Model in Real Complex Natural Environments. *Forests* **2023**, *13*, 2048. [\[CrossRef\]](#)
3. Chithambarathanum, M.; Jeyakumar, M.K. Survey on crop pest detection using deep learning and machine learning approaches. *Multimed. Tools Appl.* **2023**, *82*, 42277–42310. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Gonçalves, J.; Silva, E.; Faria, P.; Nogueira, T.; Ferreira, A.; Carlos, C.; Rosado, L. Edge-Compatible Deep Learning Models for Detection of Pest Outbreaks in Viticulture. *Agronomy* **2022**, *12*, 3052. [\[CrossRef\]](#)
5. Matsuura, T.; Yamamoto, T.; Matsumoto, Y.; Itino, T. Evolutionary host shifts across plant orders despite high host specificity in tree stem surface-living Stomaphis aphids inferred from molecular phylogeny. *J. Asia-Pac. Entomol.* **2023**, *26*, 102138. [\[CrossRef\]](#)
6. Kim, H.; Kim, D. Deep-Learning-Based Strawberry Leaf Pest Classification for Sustainable Smart Farms. *Sustainability* **2023**, *15*, 7931. [\[CrossRef\]](#)
7. Quach, L.D.; Nguyen, Q.K.; Nguyen, Q.A.; Lan, L.T.T. Rice pest dataset supports the construction of smart farming systems. *Data Brief* **2024**, *52*, 110046. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Nawaz, M.; Nazir, T.; Javed, A.; Masood, M.; Rashid, J.; Kim, J.; Hussain, A. A robust deep learning approach for tomato plant leaf disease localization and classification. *Sci. Rep.* **2022**, *12*, 18568. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Wang, R.J.; Liu, L.; Xie, C.J.; Yang, P.; Li, R.; Zhou, M. AgriPest: A Large-Scale Domain-Specific Benchmark Dataset for Practical Agricultural Pest Detection in the Wild. *Sensors* **2021**, *21*, 1601. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Akin, M.; Eydurán, S.P.; Eydurán, E.; Reed, B.M. Analysis of macro nutrient related growth responses using multivariate adaptive regression splines. *Plant Cell Tissue Organ Cult. (PCTOC)* **2020**, *140*, 661–670. [\[CrossRef\]](#)
11. Li, Z.; Jiang, X.; Jia, X.; Duan, X.; Wang, Y.; Mu, J. Classification Method of Significant Rice Pests Based on Deep Learning. *Agronomy* **2022**, *12*, 2096. [\[CrossRef\]](#)
12. Xiang, Q.C.; Huang, X.N.; Huang, Z.X.; Chen, X.M.; Cheng, J.T.; Tang, X.Y. Yolo-Pest: An Insect Pest Object Detection Algorithm via CAC3 Module. *Sensors* **2023**, *23*, 3221. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Khalid, S.; Oqaibi, H.M.; Aqib, M.; Hafeez, Y. Small Pests Detection in Field Crops Using Deep Learning Object Detection. *Sustainability* **2023**, *15*, 6815. [\[CrossRef\]](#)
14. Yuan, Z.P.; Li, S.B.; Yang, P.; Li, Y. Lightweight Object Detection Model with Data Augmentation for Tiny Pest Detection. In Proceedings of the 2022 IEEE 20th International Conference on Industrial Informatics (INDIN), Perth, Australia, 25–28 July 2022; pp. 233–238.
15. Gui, J.S.; Xu, H.R.; Fei, J.Y. Non-Destructive Detection of Soybean Pest Based on Hyperspectral Image and Attention-ResNet Meta-Learning Model. *Sensors* **2023**, *23*, 678. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Hassan, S.M.; Maji, A.K. Pest Identification Based on Fusion of Self-Attention with ResNet. *IEEE Access* **2024**, *12*, 6036–6050. [\[CrossRef\]](#)
17. Zhang, Y.R. IoT Agricultural Pest Identification Based on Multiple Convolutional Models. *J. Intent Technol.* **2023**, *24*, 905–913. [\[CrossRef\]](#)
18. Rahman, C.M.; Arko, P.S.; Ali, M.E.; Khan, M.A.I.; Apon, S.H.; Nowrin, F.; Wasif, A. Identification and Recognition of Rice Diseases and Pests Using Convolutional Neural Networks. *Biosyst. Eng.* **2020**, *194*, 112–120. [\[CrossRef\]](#)
19. Zhang, Y.E.; Liu, Y.P. Identification of navel orange diseases and pests based on the fusion of DenseNet and Self-Attention mechanism. *Comput. Intell. Neurosci.* **2021**, *2021*, 5436729. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Li, M.Z.; Cheng, S.Y.; Cui, J.Y.; Li, C.X.; Li, Z.Y.; Zhou, C.; Lv, C.L. High-Performance Plant Pest and Disease Detection Based on Model Ensemble with Inception Module and Cluster Algorithm. *Plants* **2023**, *12*, 200. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Wei, Q.F.; Li, H.; Luo, C.S.; Yu, J.; Zheng, Y.M.; Wang, F.R.; Zhang, B. Small Sample and Efficient Crop Pest Recognition Method Based on Transfer Learning and Data Transformation. *J. Comput. Methods Sci. Eng.* **2022**, *22*, 1697–1709. [\[CrossRef\]](#)
22. Tetila, E.C.; Machado, B.B.; Astolfi, G.; Belete, N.A.D.; Amorim, W.P.; Roel, A.R.; Pistori, H. Detection and Classification of Soybean Pests using Deep Learning with UAV Images. *Comput. Electron. Agric.* **2020**, *179*, 105836. [\[CrossRef\]](#)
23. Liu, J.; Wang, X. Tomato diseases and pests detection based on improved yolo v3 convolutional neural network. *Front. Plant Sci.* **2020**, *11*, 898. [\[CrossRef\]](#) [\[PubMed\]](#)

24. Sun, Y.; Liu, X.; Yuan, M.; Ren, L.; Wang, J.; Chen, Z. Automatic in-trap pest detection using deep learning for pheromone based dendroctonus valens monitoring. *Biosyst. Eng.* **2018**, *176*, 140–150. [\[CrossRef\]](#)
25. Li, Y.; Bi, W.; Jia, Y.; Wang, B.; Jin, W.; Fu, G.; Fu, X. Research on Rapid Detection for TOC in Water Based on UV-VIS Spectroscopy and 1D-SE-Inception Networks. *Water* **2023**, *15*, 2537. [\[CrossRef\]](#)
26. Yang, Z.; Zhu, L.; Wu, Y.; Yang, Y. Gated Channel Transformation for Visual Recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020), Seattle, WA, USA, 13–19 June 2020; pp. 11794–11803.
27. Qin, B.; Sun, F.; Shen, W.; Dong, B.; Ma, S.; Huo, X.; Lan, P. Deep Learning-Based Pine Nematode Trees' Identification Using Multispectral and Visible UAV Imagery. *Drones* **2023**, *7*, 183. [\[CrossRef\]](#)
28. Boniecki, P.; Zaborowicz, M.; Pilarska, A.; Piekarska-Boniecka, H. Identification Process of Selected Graphic Features Apple Tree Pests by Neural Models Type MLP, RBF and DNN. *Agriculture* **2020**, *10*, 218. [\[CrossRef\]](#)
29. Liu, D.; Lv, F.; Guo, J.; Zhang, H.; Zhu, L. Detection of Forestry Pests Based on Improved YOLOv5 and Transfer Learning. *Forests* **2023**, *14*, 1484. [\[CrossRef\]](#)
30. Wu, X.P.; Zhan, C.; Lai, Y.K.; Cheng, M.M.; Yang, J.F. IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019), Long Beach, CA, USA, 15–20 June 2019; pp. 8779–8788.
31. Peng, Y.F.; Deng, J.N.; Wang, G. Remote sensing image data enhancement based on improved SinGAN. *Chin. J. Liq. Cryst. Disp.* **2023**, *38*, 387–396. [\[CrossRef\]](#)
32. Chi, X.R.; Huang, S.Q.; Li, J.Y. Handwriting Recognition Based on Resnet-18. In Proceedings of the 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE 2021), Zhuhai, China, 24–26 September 2021; pp. 456–459.
33. Wang, Z.H.; Hu, Z.G.; Xiao, X. Crack Detection of Brown Rice Kernel Based on Optimized ResNet-18 Network. *IEEE Access* **2023**, *11*, 140701–140709. [\[CrossRef\]](#)
34. Zhao, Y.; Zhang, X.C.; Feng, W.M.; Xu, J.H. Deep Learning Classification by ResNet-18 Based on the Real Spectral Dataset from Multispectral Remote Sensing Images. *Remote Sens.* **2022**, *19*, 4883. [\[CrossRef\]](#)
35. Jin, X.; Tang, L.M.; Li, R.S.; Ji, J.T.; Liu, J. Selective Transplantation Method of Leafy Vegetable Seedlings Based on ResNet 18 Network. *Front. Plant Sci.* **2022**, *13*, 893357. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Yang, H.L.; Liu, Y.H.; Xia, T. Defect Detection Scheme of Pins for Aviation Connectors Based on Image Segmentation and Improved RESNET-50. *Int. J. Image Graph.* **2023**, *24*, 2450011. [\[CrossRef\]](#)
37. Li, C.; Cui, H.; Tian, X. A Novel CA-RegNet Model for Macau Wetlands Auto Segmentation Based on GF-2 Remote Sensing Images. *Appl. Sci.* **2023**, *13*, 12178. [\[CrossRef\]](#)
38. Yue, C.; Ye, M.Q.; Wang, P.P.; Huang, D.B.; Lu, X.J. Generative Adversarial Network Combined with SE-ResNet and Dilated Inception Block for Segmenting Retinal Vessels. *Comput. Intell. Neurosci.* **2022**, *2022*, 3585506. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Ghali, R.; Akhloufi, M.A. CT-Fire: A CNN-Transformer for Wildfire Classification on Ground and Aerial Images. *Int. J. Remote Sens.* **2023**, *44*, 7390–7415. [\[CrossRef\]](#)
40. Thomas, A.; Harikrishnan, P.M.; Palanisamy, P.; Gopi, V.P. Moving Vehicle Candidate Recognition and Classification Using Inception-ResNet-v2. In Proceedings of the 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 13–17 July 2020; pp. 467–472.
41. Wang, L.J.; Wang, J.Y.; Liu, Z.Z.; Zhu, J.; Qin, F. Evaluation of a Deep-Learning Model for Multispectral Remote Sensing of Land Use and Crop Classification. *Crop J.* **2022**, *10*, 1435–1451. [\[CrossRef\]](#)
42. He, J.; Jiang, D. Fully Automatic Model Based on SE-ResNet for Bone Age Assessment. *IEEE Access* **2021**, *9*, 62460–62466. [\[CrossRef\]](#)
43. Das, B.; Saha, A.; Mukhopadhyay, S. Rain Removal from a Single Image Using Refined Inception ResNet v2. *Circuits Syst. Signal Process.* **2023**, *42*, 3485–3508. [\[CrossRef\]](#)
44. Adbelrahman, A.; Viriri, S. FPN-SE-ResNet Model for Accurate Diagnosis of Kidney Tumors Using CT Images. *Appl. Sci.* **2023**, *17*, 9802. [\[CrossRef\]](#)
45. Xia, X.L.; Xu, C.; Nan, B. Inception-v3 for Flower Classification. In Proceedings of the 2nd International Conference on Image, Vision and Computing (ICIVC 2017), Chengdu, China, 2–4 June 2017; pp. 783–787.
46. Yang, Z.; Feng, H.; Ruan, Y.; Weng, X. Tea Tree Pest Detection Algorithm Based on Improved Yolov7-Tiny. *Agriculture* **2023**, *13*, 1031. [\[CrossRef\]](#)
47. Carranza-Flores, J.L.; Martínez-Arroyo, M.; Montero-Valverde, J.A.; Hernández-Hernández, J.L. Search for damage of the citrus miner to the lemon leaf, implementing artificial vision techniques. In *Technologies and Innovation*; International Conference on Technologies and Innovation; Springer International Publishing: Cham, Switzerland, 2020; pp. 85–97.
48. Bravo-Reyna, J.L.; Montero-Valverde, J.A.; Martínez-Arroyo, M.; Hernández-Hernández, J.L. Recognition of the damage caused by the cogollero worm to the corn plant, Using artificial vision. In *Technologies and Innovation, 6th International Conference, CITI 2020, Guayaquil, Ecuador, November 30–December 3, 2020, Proceedings*; Springer International Publishing: Cham, Switzerland, 2020; pp. 111–122.
49. Tao, Y.; He, Y.Z. Face Recognition Based on LBP Algorithm. In Proceedings of the International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, China, 25–27 September 2020; pp. 21–25.

50. Allain, B.S.; Marechal, C.; Pottier, C. Wetland Water Segmentation Using Multi-Angle and Polarimetric Radarsat-2 Datasets. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012; pp. 4915–4917.
51. Ke, Z.Y.; Ru, A.; Li, X.J. ANN based high spatial resolution remote sensing wetland classification. In Proceedings of the 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), Guiyang, China, 18–24 August 2015; pp. 180–183.
52. Gui, Y.; Li, W.; Xia, X.G.; Tao, R.; Yue, A. Infrared attention network for woodland segmentation using multispectral satellite images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5627214. [[CrossRef](#)]
53. Zhang, Y.C.; Gao, J.P.; Zhou, H.L. Breeds Classification with Deep Convolutional Neural Network. In Proceedings of the 12th International Conference on Machine Learning and Computing, Shenzhen, China, 19–21 June 2020; pp. 145–151.
54. Tang, Z.; Chen, Z.Y.; Qi, F.; Zhang, L.Y.; Chen, S.H. Pest-YOLO: Deep Image Mining and Multi-Feature Fusion for Real-Time Agriculture Pest Detection. In Proceedings of the 21st IEEE International Conference on Data Mining, Auckland, New Zealand, 7–10 December 2021; pp. 1348–1353.
55. Chen, C.J.; Huang, Y.Y.; Chen, Y.C.; Chang, C.Y.; Huang, Y.M. Identification of Fruit Tree Pests with Deep Learning on Embedded Drone to Achieve Accurate Pesticide Spraying. *IEEE Access* **2020**, *9*, 21986–21997. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.