



# Automated Identification of Insect Pests: A Deep Transfer Learning Approach Using ResNet

Christine Dewi<sup>1</sup>, Henoch Juli Christanto<sup>2\*</sup>, Guowei Dai<sup>3</sup>

<sup>1</sup> Department of Information Technology, Satya Wacana Christian University, 50715 Salatiga, Indonesia

<sup>2</sup> Department of Information System, Atma Jaya Catholic University of Indonesia, 12930 Jakarta, Indonesia

<sup>3</sup> Agricultural Information Institute of CAAS, National Agriculture Science Data Center, 100081 Beijing, China

\* Correspondence: Henoch Juli Christanto ([henoch.christanto@atmajaya.ac.id](mailto:henoch.christanto@atmajaya.ac.id))

**Received:** 09-02-2023

**Revised:** 10-24-2023

**Accepted:** 11-01-2023

**Citation:** C. Dewi, H. J. Christanto, and G. W. Dai, “Automated identification of insect pests: A deep transfer learning approach using ResNet,” *Acadlore Trans. Mach. Learn.*, vol. 2, no. 4, pp. 194–203, 2023. <https://doi.org/10.56578/ataiml020402>.



© 2023 by the authors. Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

**Abstract:** In the realm of agriculture, crop yields of fundamental cereals such as rice, wheat, maize, soybeans, and sugarcane are adversely impacted by insect pest invasions, leading to significant reductions in agricultural output. Traditional manual identification of these pests is labor-intensive and time-consuming, underscoring the necessity for an automated early detection and classification system. Recent advancements in machine learning, particularly deep learning, have provided robust methodologies for the classification and detection of a diverse array of insect infestations in crop fields. However, inaccuracies in pest classification could inadvertently precipitate the use of inappropriate pesticides, further endangering both agricultural yields and the surrounding ecosystems. In light of this, the efficacy of nine distinct pre-trained deep learning algorithms was evaluated to discern their capability in the accurate detection and classification of insect pests. This assessment utilized two prevalent datasets, comprising ten pest classes of varied sizes. Among the transfer learning techniques scrutinized, adaptations of ResNet-50 and ResNet-101 were deployed. It was observed that ResNet-50, when employed in a transfer learning paradigm, achieved an exemplary classification accuracy of 99.40% in the detection of agricultural pests. Such a high level of precision represents a significant advancement in the field of precision agriculture.

**Keywords:** Convolutional neural networks; Insect classification; Deep learning; Machine learning; Insect pests; Transfer learning

## 1 Introduction

In response to the projected escalation in global population to an estimated 10 billion by the mid-21<sup>st</sup> century, it has been recognized that a substantial increase in food production is imperative [1–3]. Crop afflictions caused by various pests lead to diseases and destruction, which in turn result in diminished yields [1, 4]. The deployment of chemical pesticides has been the conventional response to pest management, albeit with detrimental consequences for human health and environmental integrity. Consequently, agricultural scientists globally have advocated for integrated pest management (IPM) strategies, a testament to the collaborative effort in seeking alternatives to chemical pesticides [5, 6]. Pest recognition systems, underpinned by computer vision and machine learning, have emerged as pivotal in the early detection and management of pests within agricultural, forestry, and wider ecological domains. These systems are instrumental in enabling timely interventions, thus ameliorating pest-induced damages. The role of pest recognition systems is crucial in fostering sustainable agricultural practices and in the stewardship of ecosystems, as they facilitate more precise and environmentally considerate pest control measures essential for food security and the reduction of pesticide reliance.

Research into pest recognition systems is burgeoning, propelled by the imperative for sustainable pest management solutions. This research encompasses the development of sophisticated machine learning algorithms and the implementation of sensor technologies. The traditional method of manual identification is fraught with issues such as human error, subjectivity, inefficiency, limited scalability, bias, inconsistency, high costs, and intensive resource requirements. These challenges underscore the necessity for automated processes, leveraging advancements in

machine learning, computer vision, and natural language processing to augment efficacy, reduce error margins, and manage voluminous data sets with enhanced proficiency.

A myriad of autonomous pest recognition systems leveraging machine learning algorithms have been developed [3, 7]. In a notable study, Wu et al. [8] compiled an extensive insect dataset, IP102, which encompasses 102 classes and 75,000 images, and their evaluation employed various machine learning and deep learning methodologies. Concurrently, Turkoglu et al. [9] presented a multi-modal approach, both individual and hybrid, tailored for the detection of diseases and pests in apple plants, utilizing an integrated model combining a pre-trained convolutional neural network with Long Short-Term Memory (LSTM) networks [10, 11].

In related efforts, Kasinathan et al. [12] introduced a machine learning-based insect identification and classification method, exploiting Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN). This approach involved feature extraction from the Wang and Xie datasets and incorporated the GrabCut algorithm for insect detection, thus bypassing the need for manual extraction.

The utilization of transfer learning methods, where deep learning models pre-trained on extensive datasets are fine-tuned to smaller, target datasets, is well documented [13]. This technique affords several advantages, such as time and resource savings, and improved performance, particularly when labeled data for the new task is scarce [14]. Among the plethora of transfer learning strategies, the selection is contingent upon the specifics of the task and the pre-trained models available.

This study proposes an automated method for insect classification into ten categories without necessitating manual segmentation or feature extraction from images in the pre-processing phase, thereby enhancing automation compared to some extant methods [15, 16].

The primary contributions of this study are articulated as follows: Firstly, a novel system utilizing deep learning for the automated identification and classification of insects into ten categories is presented. Secondly, an analysis and validation of the transfer learning concept are provided, specifically for the ResNet model.

The subsequent structure of this document is as follows: Section 2 elucidates the proposed methodology and the deep neural networks engaged in this study. Section 3 delineates the experimental procedures, datasets employed, and the results thereof. Concluding remarks are presented in Section 4.

## **2 Material and Methods**

### **2.1 Pest Recognition Systems and Dataset**

In the realm of agricultural, forestry, and environmental management, the deployment of pest recognition systems is instrumental. These systems harness advanced technological paradigms to accurately identify and manage pest populations, thus affording a plethora of advantages to agriculturalists, land stewards, and ecosystems. The ensuing are salient benefits derived from the application of pest recognition systems:

(1) Early detection capabilities facilitate the identification of pests at incipient stages, often antecedent to observable damage to flora or ecosystems. Such preemptive discernment permits timely interventions, mitigating the propagation of pests and curtailing the magnitude of potential harm [17].

(2) Precision in pest control is augmented through the exactitude provided by these systems, enabling the implementation of pest management measures with greater specificity. Consequently, this obviates the need for indiscriminate pesticide applications, fostering an environmentally sustainable approach to pest control [15].

(3) The enhanced accuracy of pest detection and subsequent interventions culminate in the reduction of pesticide utilisation [18]. This has a cascade of benefits, encompassing monetary savings for agricultural practitioners, a diminution of environmental contaminants, and the safeguarding of agricultural workers and consumers [19].

(4) The efficacy of pest recognition systems is instrumental in safeguarding crop yields by mitigating pest-induced detriments. This safeguarding is pivotal for the sustenance of food security and the economic resilience of agricultural pursuits.

(5) The alignment of pest recognition systems with sustainable agricultural principles is marked, promoting practices that uphold environmental stewardship and economic sustainability over the long term.

(6) The systems proffer data-driven insights, which serve as a fulcrum for informed decision-making in pest management strategies [20].

(7) The optimization of pest control measures facilitated by these systems translates into substantial cost efficiencies for agriculturalists and landowners alike [21].

(8) Resistance to pesticides is a growing concern, and targeted pest control strategies made viable through these systems can play a significant role in managing this resistance [22].

(9) The health of ecosystems, particularly within the domains of forestry and natural resource management, is bolstered through the mitigation of invasive species and detrimental pests, contributing to ecological resilience.

(10) Remote monitoring capabilities inherent in some pest recognition systems, through the integration of drones or IoT devices, afford expansive area surveillance with minimal manual intervention [23].

(11) Adaptability and scalability are intrinsic to these systems, enabling their application across diverse pest types and geographic locales. This flexibility is essential in the face of emergent pest threats or evolutionary changes within extant populations.

(12) Centralization and dissemination of pest-related data facilitate collaborative pest management endeavours, unifying researchers, agriculturists, and consortia in the pursuit of efficacious control strategies [24].

Convolutional Neural Networks (CNNs) have emerged as a potent tool in the identification and classification of pests, with proven efficacy across agricultural, forestry, and environmental monitoring applications. The intrinsic design of CNNs renders them particularly adept at discerning and extrapolating salient features from image-based data, thereby serving as a cornerstone technology in pest recognition tasks [25].

The dataset proposed by Deng et al. [26], comprising ten disparate pest categories depicted through photographic images, has been explored and harnessed within these studies [27, 28]. Visual exemplars from each pest category are depicted in Figure 1, while Table 1 delineates the categorical data of the 'Small Dataset' as utilised for pest classification.



**Figure 1.** Representative images from dataset

**Table 1.** Summary of 'Small Dataset' (Deng et al. [26] data) for pest classification

Class	Number of Images
<i>Locusta migratoria</i>	72
<i>Parasa lepida</i>	59
<i>Chrysochus chinensis</i>	46
<i>Spodoptera exigua</i>	56
<i>Atractomorpha sinensis</i>	62
<i>Empoasca flavescens</i>	41
<i>Laspeyresia pomonella</i>	65
<i>Spodoptera exigua</i> larva	68
Gypsy moth larva	40
<i>Laspeyresia pomonella</i> larva	56
Total number of images	565

## 2.2 Proposed Methodology

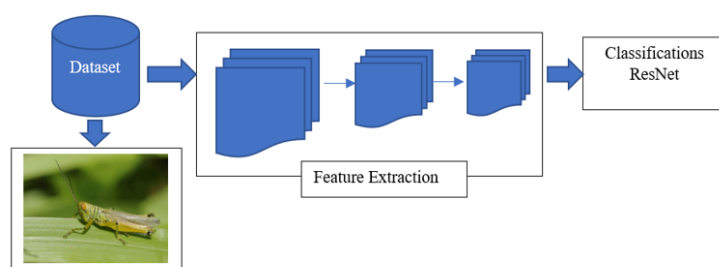
The application of transfer learning to the classification of entomological imagery is elucidated in Figure 2. Within this framework, the latent convolutional layers of pre-existing neural networks are meticulously adjusted to facilitate a categorization into ten distinct classes. It is posited that this nuanced calibration of transfer learning, effected through the modification of terminal layers within the network, augments the network's discriminative power, thereby enhancing performance.

Transfer learning emerges as a pivotal technique in the development of pest recognition systems, particularly in scenarios marked by a scarcity of labeled data specific to the task at hand. The process is delineated as follows: Initially, the task is defined with precision. Subsequently, a collection of labeled images embodying the pests of interest is compiled, with annotations encompassing both the presence and absence of pests, and, where pertinent, the exact species. This dataset is anticipated to be representative, encompassing various life stages and environmental contexts [29, 30]. A model with prior training on a comprehensive dataset, such as ResNet 50 or ResNet101 utilizing the ImageNet database, is selected. The data are then subjected to preparatory processing, which involves resizing,

normalizing, and augmenting the images to conform to the expected input schema of the chosen model. Techniques for data augmentation may include, but are not limited to, random cropping, rotation, and flipping [31].

Alterations are made to the pre-trained model to tailor it to the specific task of pest recognition. The model is then trained on the labeled pest dataset, applying transfer learning principles by initializing the model with pre-trained weights and refining these on the dataset in question [32]. Monitoring of the training sequence is conducted, employing strategies such as early stopping to mitigate overfitting, with hyperparameters being adjusted accordingly. Model performance is evaluated on a distinct validation set, with hyperparameters fine-tuned to optimize outcomes based on the validation feedback. A comprehensive assessment of the trained model is performed using a dedicated test dataset to determine the model's generalization capability. Metrics of accuracy and reliability are computed to gauge model performance [33, 34]. Should the model satisfy predefined performance thresholds, it is then deployed for real-time or batch inference within the pest recognition framework, potentially integrating with remote sensing technologies such as drones or stationary cameras.

The principle of transfer learning leverages the knowledge acquired from expansive image datasets by pre-trained models, applying it to the specialized domain of pest detection and classification [35]. This approach holds the potential to markedly diminish the volume of labeled data required to construct an efficacious pest recognition system.



**Figure 2.** Application of transfer learning to the classification of insect pests

In the study presented, transfer learning has been utilized within deep neural networks to facilitate the classification of insect pests. Transfer learning is characterized as a strategy whereby knowledge harnessed by a pre-established model is repurposed to solve analogous problems. The training of deep neural networks traditionally necessitates extensive datasets, considerable computational time, and processing power.

It is posited that the integration of transfer learning, particularly with pre-trained deep neural networks, constitutes an advanced, cost-effective alternative for classification challenges in scenarios marked by data and processing resource scarcity. The selection of an appropriate pre-trained neural network model is pivotal within the transfer learning paradigm. This critical choice is contingent upon the similarity of the existing problem to the task for which the model was initially trained. A heightened risk of overfitting is posited when the source dataset employed for training exhibits significant congruence in size and composition with the target dataset.

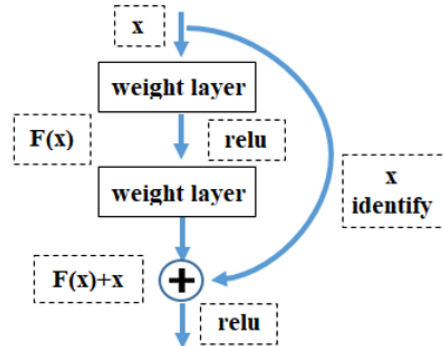
### 2.3 ResNet

Within the domain of convolutional neural network architectures, the ResNet family stands as a prominent contribution, as introduced by He et al. [36]. The moniker "ResNet" is derived from "Residual Network," a term that encapsulates the network's pioneering innovation: the integration of residual blocks. These blocks are instrumental in enabling the training of profoundly deep neural networks. The ResNet architecture is presented in two variants, namely ResNet-50 and ResNet-101, the numerals indicative of the layers' depth [36]. Both models function akin to standard deep networks while additionally featuring identity mapping capabilities. It is through this innovation that ResNet architectures address and mitigate the issue of vanishing gradients, with ResNet-50 boasting an architecture that scales from 34 to an expansive 152 layers [37, 38]. Figure 3 delineates the ILSVRC 2015 classification task's winning architecture, a testament to the architecture's efficacy [39].

The deeper variant, ResNet-101, extends to 101 layers, incorporating an increased number of residual blocks to capture more nuanced patterns and features within the dataset. Its application is deemed appropriate when tasks necessitate heightened accuracy or intricate feature extraction, such as image classification and object recognition. Inherently, the added depth translates to a greater parameter count, with ResNet-101 possessing approximately 44.5 million parameters, rendering it a more potent albeit computationally demanding model.

The schematic of a residual block is depicted in Figure 3. It is observed that for applications such as fault diagnosis, ResNet-50, conceptualized as a TCNN, is employed, leveraging its strength as a feature extractor post-training on the ImageNet dataset [40]. The amalgamation of convolutional neural networks with transfer learning is

demonstrated as a method for fault diagnostics. Moreover, the predictive capacity of ResNet-50 in the early diagnosis of Alzheimer's disease (AD) has been explored [41]. Utilizing magnetic resonance imaging (MRI) scans, ResNet-50 facilitates multi-class classification to ascertain the presence and intensity of clinical dementia ratings (CDR). This machine learning approach has demonstrated high precision in classifying AD, suggesting potential utility in the preemptive identification of AD patients prior to clinical evaluation.



**Figure 3.** Residual block

The ResNet-101 architecture represents an advancement in the field of deep convolutional neural networks, encompassing a comprehensive structure that addresses a multitude of computer vision tasks. Characterised by its depth, with an assemblage of 101 layers, ResNet-101 extends beyond its predecessor, ResNet-50, by facilitating enhanced feature extraction through additional layers. The architecture is delineated as follows: Initially, an input layer receives an image comprising typically three colour channels, such as red, green, and blue (RGB). Subsequent to the input layer, a series of convolutional layers are employed, applying a multitude of convolutional filters to extract low-level and mid-level features, encompassing edges, textures, and shapes. At the heart of ResNet-101 lies the implementation of residual blocks, which are pivotal in mitigating the vanishing gradient problem, thereby enabling the training of deep networks. These blocks consist of convolutional layers, batch normalization, and ReLU activation functions, with the addition of skip connections that perform elementwise addition of the input from a preceding layer to the output of the block. Furthermore, ResNet-101 incorporates two distinct types of residual blocks: identity blocks, which maintain input and output dimensions, and projection blocks, which modify dimensions via a  $1 \times 1$  convolutional layer. Pooling layers are interspersed following the residual blocks, typically adopting average pooling to diminish spatial dimensions while preserving salient features. As the network progresses towards its conclusion, fully connected layers are utilised for classification, culminating in an output layer where a softmax activation function articulates class probabilities for multi-class classification tasks.

The ResNet-101 architecture has been demonstrated to be a flexible construct, extensively applied across diverse computer vision domains. Its utility spans several key applications as follows:

(1) Image classification. ResNet-101 is often selected for image classification endeavors. Its architectural depth enables the intricate capture of patterns within images, a process critical for accurate categorization into predefined classes.

(2) Object detection. The capacity of ResNet-101 to serve as a potent feature extractor has been validated in object detection frameworks such as Faster R-CNN and Mask R-CNN. Here, it aids in the pinpointing and classification of objects across still images and video sequences.

(3) Semantic segmentation. The architecture is actively employed in semantic segmentation tasks, with its depth providing the means to execute detailed classification of individual pixels into specified categories.

(4) Transfer learning: The adaptability of ResNet-101 is particularly evident in transfer learning scenarios. Models pretrained on comprehensive datasets like ImageNet present an invaluable foundation for a multitude of vision-based projects, where fine-tuning on smaller datasets can elicit commendable outcomes despite limited sample volumes.

(5) Fine-grained classification. Tasks requiring discernment between closely related categories within a larger class benefit from the depth of ResNet-101, which facilitates nuanced differentiation.

(6) Medical image analysis. Its application has been extended to medical image analysis, addressing critical tasks such as disease detection, organ segmentation, and pathology classification.

(7) Scene recognition. Employed in the recognition and classification of scenes within images, ResNet-101 supports applications in autonomous navigation and content tagging.

In sum, the depth and structural ingenuity of ResNet-101, featuring residual blocks and skip connections, render it a robust mechanism for complex feature extraction challenges. This has solidified its status as a cornerstone in the



computer vision research community and operational domain.

### 3 Result and Discussion

In the current investigation, the efficacy of transfer learning in classifying insect pest images from the Deng dataset was evaluated across nine distinct deep neural network architectures. Transfer learning, coupled with fine-tuning strategies, was employed to mitigate the risk of overfitting, a frequent challenge in machine learning tasks with limited data. Examination of the classification outcomes, as detailed in Table 2, indicates that with a diminutive dataset, the ResNet-50 model exhibited superior performance, achieving an accuracy of 99.40% and a precision of 99.10%. In comparison, the ResNet-101 model attained an accuracy of 97.63% and a precision of 97.69%.

**Table 2.** Detailed classification results

Model	Accuracy	Precision	Recall	F1-score	Sensitivity	Specificity
ResNet-50	99.40 %	99.10 %	99.10 %	99.10 %	99 %	99.10 %
ResNet-101	97.63 %	97.69 %	97.40 %	97.75 %	97 %	97.50 %

Confusion Matrix											
Output Class	Cydia pomonella	121 7.0%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	97.6% 2.4%
	Gryllotalpa	0 0.0%	152 8.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	Oriental fruit fly	0 0.0%	0 0.0%	139 8.0%	0 0.0%	0 0.0%	3 0.2%	0 0.0%	0 0.0%	0 0.0%	97.9% 2.1%
	Pieris rapae Linnaeus	0 0.0%	0 0.0%	1 0.1%	170 9.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.4% 0.6%
	Spodoptera litura	0 0.0%	0 0.0%	0 0.0%	0 0.0%	129 7.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	Weevil	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	160 9.2%	1 0.1%	0 0.0%	0 0.3%	95.8% 4.2%
	leafhopper	1 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	124 7.2%	0 0.0%	0 0.0%	99.2% 0.8%
	locust	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	220 12.7%	0 0.0%	1 0.1%	99.1% 0.9%
	snail	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	316 18.2%	0 0.0%	100% 0.0%
	stinkbug	2 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	1 0.1%	0 0.0%	0 0.0%	97.3% 2.7%
		97.6% 2.4%	100% 0.0%	99.3% 0.7%	100% 0.0%	98.5% 1.5%	97.0% 3.0%	97.6% 2.4%	99.5% 0.5%	100% 0.0%	96.2% 3.8%

**Figure 4.** Confusion matrix for the highest-performing neural network (ResNet-50)

The computational environment deployed for model training comprised an Nvidia GTX3060 Super GPU accelerator and an AMD Ryzen 7 3700X CPU, featuring an 8-core processor with 32GB of DDR4-3200 memory. A confusion matrix for the highest-performing neural network, ResNet-50, is illustrated in Figure 4. For comprehensive performance evaluation of the deep neural networks, metrics including accuracy, precision, recall, F1-score, sensitivity, and specificity were calculated. These metrics offer a multi-dimensional perspective on model efficacy, facilitating a nuanced analysis of the nine transfer learning models. Eqs. (1)-(4) provide the computational framework for these metrics, which are integral to validating the performance of the aforementioned models [32]. Classification accuracy is defined as the ratio of correct predictions to the total number of datasets.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

Precision, denoted as positive predictive values, reflects the proportion of true positive predictions in all positive classes.

$$\text{Precision} = TP / (FP + TP) \quad (2)$$

Recall measures the fraction of true positives identified out of all actual positives.

$$\text{Recall} = FN / (FN + TP) \quad (3)$$

The F1-score represents the harmonic mean of precision and recall.

$$F1 - \text{Score} = 2 * \frac{(\text{Precision} \cdot \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (4)$$

In this section, sensitivity and specificity were meticulously calculated. Sensitivity, or recall, refers to the proportion of correctly identified positive instances relative to the total number of actual positive instances. A perfect sensitivity score is 1.0, indicating flawless prediction, whereas a score of 0.0 signifies complete misclassification. Specificity, similarly, is determined by the ratio of true negative predictions to the total number of negative instances, with 1.0 being the ideal and 0.0 the antithesis of accurate prediction.

**Table 3.** Comparative analysis with preceding state-of-the-art research

Work	Dataset	Method	Accuracy
Nanni et al. [42]	Small (Deng et al. [26])	CNN	95.52 %
Deng et al. [26]	Small (Deng et al. [26])	LCP+SVM	85.5 %
Wang et al. [43]	Small (Deng et al. [26])	TL AlexNet	93.84 %
Proposed Method	Small (Deng et al. [26])	ResNet-50	99.40 %

Table 3 describes the comparative analysis with preceding state-of-the-art research. When juxtaposed with preceding benchmarks in the field, the method delineated herein outstripped the performance of the counterparts, achieving an accuracy pinnacle of 99.40%. This is followed by the work of Nanni et al. [42], which attained a commendable 95.52% accuracy. The adoption of ResNet models, particularly ResNet-50, in deep learning and computer vision has been marked by numerous advantages. A synthesis of the merits reveals the following:

(1) Facilitation of deep network training. ResNet models incorporate skip connections, or residual blocks, which have been shown to alleviate the vanishing gradient issue. This architectural feature permits the training of networks with an extensive number of layers, thus averting performance degradation despite increased depth.

(2) Enhanced training efficiency. It has been observed that ResNet architectures often exhibit accelerated convergence during the training phase when contrasted with their predecessors. The skip connections facilitate easier learning of identity mappings, culminating in a reduced number of iterations required to reach optimal performance.

(3) Gradient flow optimisation. The integration of skip connections in ResNet models underpins a robust gradient signal across the network. Such an enhancement in gradient flow contributes to rapid convergence and stabilises the training of deeply layered networks.

(4) Benchmark performance. Models such as ResNet-101 and ResNet-152 have consistently delivered exemplary performance on a variety of computer vision benchmarks. Their proficiency encompasses tasks ranging from image classification to semantic segmentation.

(5) Versatile application. The utility of ResNet models spans a broad spectrum of computer vision tasks, evidencing their adaptability to varying challenges including but not limited to classification, detection, and segmentation. This versatility extends to transfer learning applications, where models pre-trained on extensive datasets prove beneficial.

(6) Potential of transfer learning. The availability of pre-trained ResNet models on datasets such as ImageNet facilitates their use in transfer learning contexts. Such utilization allows for a reduction in training duration while maintaining commendable results in scenarios characterised by limited data.

(7) Generalisation capability. Notably, ResNet models have demonstrated an impressive capacity to generalise across diverse datasets and contexts, affirming their applicability in a multitude of real-world scenarios.

(8) Architectural interpretability. The architecture of ResNet models, owing to the presence of skip connections, presents an enhanced level of interpretability. Such a structure permits a more coherent analysis of information flow and feature representation across different layers.

(9) Community and resource availability. The widespread adoption of ResNet models has fostered a substantial user community and a wealth of resources, ranging from pre-trained models to extensive literature, thereby easing the integration and utilization of these models in both research and practical applications.

(10) Discriminative feature learning. Deep feature representations learned by ResNet models are distinguished by their discriminative and transferable qualities, proving valuable for downstream tasks irrespective of the initial training objective.

In the landscape of computer vision, ResNet architectures have been discerned to offer a robust and potent framework, as evidenced by their extensive deployment across varied tasks within the domain. The profundity of their architectural design, coupled with the facilitation of training, has been observed to foster strong gradient flow and perpetuate a consistent performance that aligns with the apex of contemporary standards. These attributes have been pivotal in cementing the role of ResNet models as a foundational element in the advancement of deep learning methodologies and their practical applications within the sphere of computer vision.

#### 4 Conclusions

The study delineated herein has leveraged the ResNet-50 and ResNet-101 frameworks to evaluate their efficacy in the identification, detection, and classification of insect pests. The application of ResNet in pest recognition has been substantiated as an efficacious methodology, particularly when utilizing pre-trained models in conjunction with limited labeled data sets. The objective pursued was to augment agricultural productivity by facilitating an automated, transfer learning-based mechanism for the prompt and effective detection of insect pests.

Analyses performed have revealed that the ResNet-50 transfer learning algorithm exhibits remarkable proficiency in the detection of agricultural pests, culminating in an unprecedented accuracy rate of 99.40%, thereby surpassing other transfer learning algorithms in performance metrics. This represents a significant stride in precision agriculture technologies. Prospective studies are advocated to further probe the potential of transfer learning within deep neural networks, with an emphasis on the minimization of computational time complexity inherent in the identification of insect pests. Such inquiries are anticipated to refine the efficiency and applicability of transfer learning algorithms in real-world agricultural settings.

#### Data Availability

The data used to support the research findings are available from the corresponding author upon request.

#### Acknowledgements

The author would like to thank all colleagues from Satya Wacana Christian University, Indonesia, and all involved in this research.

#### Conflicts of Interest

The authors declare no conflict of interest.

#### References

- [1] V. Malathi and M. P. Gopinath, "Classification of pest detection in paddy crop based on transfer learning approach," *Acta Agric. Scand. Sect. B Soil Plant Sci.*, vol. 71, no. 7, pp. 552–559, 2021. <http://doi.org/10.1080/09064710.2021.1874045>
- [2] T. T. Nguyen, Q. T. Vien, and H. Sellaheewa, "An efficient pest classification in smart agriculture using transfer learning," *EAI Endorsed Trans. Ind. Networks Intell. Syst.*, vol. 8, no. 26, pp. 1–8, 2021. <http://doi.org/10.4108/eai.26-1-2021.168227>
- [3] C. Dewi, R. C. Chen, Y. C. Zhuang, and H. J. Christanto, "Yolov5 series algorithm for road marking sign identification," *Big Data Cogn. Comput.*, vol. 6, no. 4, p. 149, 2022. <http://doi.org/10.3390/bdcc6040149>
- [4] C. Yu, Z. X. Liu, X. J. Liu, F. Qiao, Y. Wang, F. G. Xie, Q. Wei, and Y. Yang, "A DenseNet feature-based loop closure method for visual SLAM system," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Dali, China, 2019, pp. 258–265. <http://doi.org/10.1109/ROBIO49542.2019.8961714>
- [5] H. Kibriya, R. Amin, A. H. Alshehri, M. Masood, S. S. Alshamrani, and A. Alshehri, "A novel and effective brain tumor classification model using deep feature fusion and famous machine learning classifiers," *Comput. Intell. Neurosci.*, vol. 2022, p. 7897669, 2022. <http://doi.org/10.1155/2022/7897669>
- [6] T. Jintasuttisak, E. Edirisinghe, and A. Elbattay, "Deep neural network based date palm tree detection in drone imagery," *Comput. Electron. Agric.*, vol. 192, 2022. <http://doi.org/10.1016/j.compag.2021.106560>
- [7] C. Dewi and R. C. Chen, "Complement naive bayes classifier for sentiment analysis of internet movie database," in *Asian Conference on Intelligent Information and Database Systems*, Ho Chi Minh City, Vietnam, November 28–30, 2022, pp. 81–93. [https://doi.org/10.1007/978-3-031-21743-2\\_7](https://doi.org/10.1007/978-3-031-21743-2_7)
- [8] X. P. Wu, C. Zhan, Y. K. Lai, M. M. Cheng, and J. F. Yang, "IP102: A large-scale benchmark dataset for insect pest recognition," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, pp. 8787–8796. <https://doi.org/10.1109/CVPR.2019.00899>
- [9] M. Turkoglu, D. Hanbay, and A. Sengur, "Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests," *J. Ambient Intell. Humaniz. Comput.*, vol. 13, pp. 3335–3345, 2022. <http://doi.org/10.1007/s12652-019-01591-w>



- [10] C. Dewi, R. C. Chen, H. Yu, and X. Y. Jiang, "Robust detection method for improving small traffic sign recognition based on spatial pyramid pooling," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, pp. 1–18, 2021. <http://doi.org/10.1007/s12652-021-03584-0>
- [11] R. C. Chen, C. Dewi, S. W. Huang, and R. E. Caraka, "Selecting critical features for data classification based on machine learning methods," *J. Big Data*, vol. 7, no. 1, p. 52, 2020. <http://doi.org/10.1186/s40537-020-00327-4>
- [12] T. Kasinathan, D. Singaraju, and S. R. Uyyala, "Insect classification and detection in field crops using modern machine learning techniques," *Inf. Process. Agric.*, vol. 8, no. 3, pp. 446–457, 2021. <http://doi.org/10.1016/j.inpa.2020.09.006>
- [13] A. Rehman, S. Naz, M. I. Razzak, and I. A. Hameed, "Automatic visual features for writer identification: A deep learning approach," *IEEE Access*, vol. 7, pp. 17 149–17 157, 2019. <http://doi.org/10.1109/ACCESS.2018.2890810>
- [14] C. Dewi and R. C. Chen, "Combination of ResNet and spatial pyramid pooling for musical instrument identification," *Cybern. Inf. Technol.*, vol. 22, no. 1, pp. 104–116, 2022. <https://doi.org/10.2478/cait-2022-0007>
- [15] C. Dewi, R. C. Chen, and H. Yu, "Weight analysis for various prohibitory sign detection and recognition using deep learning," *Multimed. Tools Appl.*, vol. 79, pp. 32 897–32 915, 2020. <http://doi.org/10.1007/s11042-020-09509-x>
- [16] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, 2021. <http://doi.org/10.1080/07391102.2020.1788642>
- [17] C. J. Chen, Y. Y. Huang, Y. S. Li, Y. C. Chen, C. Y. Chang, and Y. M. Huang, "Identification of fruit tree pests with deep learning on embedded drone to achieve accurate pesticide spraying," *IEEE Access*, vol. 9, pp. 21 986–21 997, 2021. <http://doi.org/10.1109/ACCESS.2021.3056082>
- [18] R. J. Wang, R. Li, T. J. Chen, J. Zhang, C. J. Xie, K. Qiu, P. Chen, J. M. Du, H. B. Chen, F. R. Shao, H. Y. Hu, and H. Y. Liu, "An automatic system for pest recognition and forecasting," *Pest Manag. Sci.*, vol. 78, no. 2, pp. 711–721, 2022. <http://doi.org/10.1002/ps.6684>
- [19] M. E. Karar, F. Alsunaydi, S. Albusaymi, and S. Alotaibi, "A new mobile application of agricultural pests recognition using deep learning in cloud computing system," *Alexandria Eng. J.*, vol. 60, no. 55, pp. 4423–4432, 2021. <http://doi.org/10.1016/j.aej.2021.03.009>
- [20] M. Á. Rodríguez-García, F. García-Sánchez, and R. Valencia-García, "Knowledge-based system for crop pests and diseases recognition," *Electronics*, vol. 10, no. 8, p. 905, 2021. <http://doi.org/10.3390/electronics10080905>
- [21] C. Dewi and H. J. Christanto, "Automatic medical face mask recognition for COVID-19 mitigation: Utilizing YOLO V5 object detection," *Rev. d'Intelligence Artif.*, vol. 37, no. 3, pp. 627–638, 2023. <http://doi.org/10.18280/ria.370312>
- [22] L. Liu, C. J. Xie, R. J. Wang, P. Yang, S. Sudirman, J. Zhang, R. Li, and F. Y. Wang, "Deep learning based automatic multiclass wild pest monitoring approach using hybrid global and local activated features," *IEEE Trans. Ind. Informatics*, vol. 17, no. 11, pp. 7589–7598, 2021. <http://doi.org/10.1109/TII.2020.2995208>
- [23] C. Dewi, R. C. Chen, H. Yu, and X. Y. Jiang, "XAI for image captioning using SHAP," *J. Inf. Sci. Eng.*, vol. 39, no. 34, pp. 711–724, 2023. [https://doi.org/10.6688/JISE.202307\\_39\(4\).0001](https://doi.org/10.6688/JISE.202307_39(4).0001)
- [24] C. Dewi, A. P. S. Chen, and H. J. Christanto, "Deep learning for highly accurate hand recognition based on Yolov7 model," *Big Data Cogn. Comput.*, vol. 7, no. 1, p. 53, 2023. <http://doi.org/10.3390/bdcc7010053>
- [25] V. S. Saravananarajan, R. C. Chen, C. Dewi, L. S. Chen, and L. Ganesan, "Car crash detection using ensemble deep learning," *Multimed. Tools Appl.*, pp. 1–19, 2023. <http://doi.org/10.1007/s11042-023-15906-9>
- [26] L. M. Deng, Y. J. Wang, Z. Z. Han, and R. S. Yu, "Research on insect pest image detection and recognition based on bio-inspired methods," *Biosyst. Eng.*, vol. 169, pp. 139–148, 2018. <http://doi.org/10.1016/j.biosystemseng.2018.02.008>
- [27] R. Sharma, A. Singh, N. Z. Jhanjhi, M. Masud, E. S. Jaha, and S. Verma, "Plant disease diagnosis and image classification using deep learning," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 2125–2140, 2022. <http://doi.org/10.32604/cmc.2022.020017>
- [28] J. F. Gao, J. C. Westergaard, E. H. R. Sundmark, M. Bagge, E. Liljeroth, and E. Alexandersson, "Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning," *Knowledge-Based Syst.*, vol. 214, 2021. <http://doi.org/10.1016/j.knosys.2020.106723>
- [29] R. C. Chen, C. Dewi, Y. C. Zhuang, and J. K. Chen, "Contrast limited adaptive histogram equalization for recognizing road marking at night based on YOLO models," *IEEE Access*, 2023. <https://doi.org/10.1109/ACCESS.2023.3309410>
- [30] C. Dewi, R. C. Chen, and Y. T. Liu, "Synthetic traffic sign image generation applying generative adversarial networks," *Vietnam J. Comput. Sci.*, vol. 9, no. 3, pp. 333–348, 2022. <https://doi.org/10.1142/S2196888822500191>

- [31] C. Dewi, R. C. Chen, and S. K. Tai, "Evaluation of robust spatial pyramid pooling based on convolutional neural network for traffic sign recognition system," *Electronics*, vol. 9, no. 6, p. 889, 2022. <https://doi.org/10.3390/electronics9060889>
- [32] C. Dewi, R. Chen, Y. T. Liu, and H. Yu, "Various generative adversarial networks model for synthetic prohibitory sign image generation," *Appl. Sci.*, vol. 11, no. 7, p. 2913, 2021. <https://doi.org/10.3390/app11072913>
- [33] G. W. Dai, J. C. Fan, and C. Dewi, "ITF-WPI: Image and text based cross-modal feature fusion model for wolfberry pest recognition," *Comput. Electron. Agric.*, vol. 212, p. 108129, 2023. <https://doi.org/10.1016/j.compag.2023.108129>
- [34] S. Javed, A. Sajid, T. Kiren, I. U. Khan, C. Dewi, F. Cauteruccio, and H. J. Christanto, "A subjective logical framework-based trust model for wormhole attack detection and mitigation in Low-Power and Lossy (RPL) IoT-Networks," *Information*, vol. 14, no. 9, p. 478, 2023. <https://doi.org/10.3390/info14090478>
- [35] C. Dewi, A. P. S. Chen, and H. J. Christanto, "Recognizing similar musical instruments with YOLO models," *Big Data Cogn. Comput.*, vol. 7, no. 2, p. 94, 2023. <https://doi.org/10.3390/bdcc7020094>
- [36] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27-30 June, 2016, pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- [37] G. Chander, B. L. Markham, and D. L. Helder, "Summary of current radiometric calibration coefficients for landsat MSS, TM, ETM+, and EO-1 ALI sensors," *Remote Sens. Environ.*, vol. 113, no. 5, pp. 893–903, 2009. <http://doi.org/10.1016/j.rse.2009.01.007>
- [38] W. Z. Fang, C. G. Wang, X. Chen, W. Wan, H. Li, S. Y. Zhu, Y. Fang, B. J. Liu, and Y. Hong, "Recognizing global reservoirs from landsat 8 images: A deep learning approach," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 9, pp. 3168–3177, 2019. <http://doi.org/10.1109/jstars.2019.2929601>
- [39] Y. N. Ibrahim, H. B. Wang, M. Bai, Z. Liu, J. N. Wang, Z. M. Yang, and Z. M. Chen, "Soft error resilience of deep residual networks for object recognition," *IEEE Access*, vol. 8, pp. 19 490–19 503, 2020. <http://doi.org/10.1109/ACCESS.2020.2968129>
- [40] L. Wen, X. Y. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Comput. Appl.*, vol. 32, pp. 6111–6124, 2020. <http://doi.org/10.1007/s00521-019-04097-w>
- [41] L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan, and C. P. Fulton, "Classification of Alzheimer's disease with and without imagery using gradient boosted machines and ResNet-50," *Brain Sci.*, vol. 9, no. 9, p. 212, 2019. <http://doi.org/10.3390/brainsci9090212>
- [42] L. Nanni, G. Maguolo, and F. Pancino, "Insect pest image detection and recognition based on bio-inspired methods," *Ecol. Inform.*, vol. 57, p. 101089, 2020. <http://doi.org/10.1016/j.ecoinf.2020.101089>
- [43] D. W. Wang, L. M. Deng, J. G. Gao, J. Y. Gao, H. F. Zhu, and Z. Z. Han, "Recognition pest by image-based transfer learning," *J. Sci. Food Agric.*, vol. 99, no. 10, pp. 4524–4531, 2019. <http://doi.org/10.1002/jsfa.9689>