SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Modified RESNET50 with Attention Module for Detection and Classification of Pests in Vegetable Crops

R. Prabha[1,*], K. Selvan[1]

[1] Research Scholar, P.G. and Research Department of Computer Science, J.J. College of Arts and Science (Autonomous), Pudukkottai, Affiliated to Bharathidasan University, Tiruchirappalli, India
[2] Assistant Professor & Research Advisor, P.G. and Research Department of Computer Science, J.J. College of Arts and Science (Autonomous), Pudukkottai, Affiliated to Bharathidasan University, Tiruchirappalli, India

ABSTRACT

The agricultural sector plays a pivotal role in ensuring global food security. This study addresses the significant challenge of pest infestations in vegetable crops by automating pest identification and classification through deep learning techniques. We utilize a state-of-the-art Convolutional Neural Network (CNN), specifically a Modified ResNet50 architecture enhanced with an Adversarial Attention Module. This approach is designed to improve feature extraction and model performance. The purpose of our study is to develop and evaluate a model that can accurately identify and classify pest species, thereby aiding in timely pest management. The ResNet50 backbone, pre-trained on an extensive dataset of pest-crop interactions, is augmented with an attention module to refine its capabilities. Performance is evaluated on a hold-out dataset of previously unseen images, where the Modified ResNet50 achieves a classification accuracy of [insert quantitative result, e.g., 92%]. Comparative analysis shows that our model outperforms other deep learning models by [insert quantitative result, e.g., 5%] in precision and recall metrics. This research contributes to precision agriculture by offering a more effective and environmentally friendly pest identification solution, which supports improved pest management strategies and enhances global food security.

Keywords:
Pest identification and classification; deep learning techniques; convolutional neural network (CNN); modified RESNET50 attention module; data augmentation

1. Introduction

The global agricultural landscape faces a persistent challenge in ensuring food security for an ever-growing population while optimizing resource use and minimizing environmental impacts. Among the myriad threats to agricultural productivity, pest infestations in vegetable crops pose a significant and recurrent obstacle. The unchecked proliferation of pests can result in substantial yield losses, economic hardships for farmers, and, in some cases, environmental damage due to increased pesticide usage. In order to tackle this pressing problem, this research explores the field of deep learning-based insect identification and categorization in vegetable crops, providing a viable way

* *Corresponding author.*
E-mail address: prabha.pr@gmail.com

https://doi.org/10.37934/araset.63.1.6786

forward to transform pest control techniques. Conventional approaches to pest identification and control often depend on manual scouting, which is time-consuming, labour-intensive, and prone to human mistakes. Moreover, chemical pesticides are the mainstay of conventional pest management techniques, and when misused or overused, they can be harmful to human health and the environment. In light of these challenges, there is a growing urgency to develop more efficient, precise, and sustainable pest management approaches. Within the field of artificial intelligence (AI), deep learning has become a disruptive technique, especially in computer vision applications. The implementation of deep learning methods, such as convolutional neural networks (CNNs), has allowed machines to detect patterns and generate predictions with astonishing precision. This technology's potential extends to agriculture, where it can be harnessed to automate the process of pest detection and classification in vegetable crops. This study focuses on three core objectives:

1.1 Data Collection and Pre-Processing

To develop effective deep learning models, high-resolution image datasets are collected, capturing various insect species and their interactions with different vegetables. Maintaining data uniformity and quality is crucial for model success. This involves ensuring that images are consistently high-resolution, properly labelled, and representative of diverse pest behaviours and vegetable types. By achieving high data quality and consistency, the model's ability to accurately detect and classify pests is enhanced, leading to more reliable and actionable insights for pest management. Figure 1 shows the sample pest images obtained from the database.

Fig. 1. Pest images in the database

1.2 Model Development and Training

RESNET-50 is tailored specifically for pest identification and classification by leveraging advanced deep learning architectures. Transfer learning techniques are applied to fine-tune the model, enhancing its performance even with limited labelled data. This approach ensures that the model effectively learns and adapts to the nuances of pest features and classifications, resulting in improved accuracy and efficiency in pest detection and management.

1.3 Real-Time Implementation

In addition to developing the deep learning model, practical application within agricultural settings is a key focus. This involves creating user-friendly interfaces and integrating the system with existing agricultural technologies, such as drones and autonomous vehicles, for real-time pest monitoring and management. The research outcomes promise substantial benefits for the agricultural sector and broader sustainability efforts. By automating pest management, the system aims to significantly reduce crop losses, decrease reliance on chemical pesticides, and promote more environmentally friendly and efficient agricultural practices. This study marks a crucial advancement in precision agriculture, enhancing pest management in vegetable crop production and supporting sustainable farming practices.

2. Related Work

Many scientists have proposed various machine learning models for pest detection and similar applications some of them are. Li, Zhen and Li [1] proposed a residual network that provide an accuracy of 70% using a convolutional neural network. Ramachandran& Sathurshan [2] worked mainly on urban pests using MobileNetV3 and obtained a commendable improvement in accuracy with a dataset of 300 images of 10 classes of pests in urban areas. Osco-Mamani and Israel [3] particularly focused on the Olive leaf disease using deep learning architectures. Li *et al.,* [4] review proposed a smart pest monitoring systems using Deep learning methods. Such methods are particularly focused on single or similar varieties of pest species. Liu *et al.,* [5] created a novel network called "PestNet" for automated multi-class pest detection and has achieved an accuracy of 75.46% on Multi-class Pest Dataset 2018 (MPD2018). This network makes use of CNN and RPN with a backbone of contextual. The work of Rustia *et al.,* [6] aims to improve the efficiency of pest control, describes a cascaded deep learning classification technique for automatically identifying and detecting insect pests in greenhouses it considers 4 classes of pests invading the greenhouses. The methodology, which attempts to improve the accuracy to 90% is presented in the publication by Türkoğlu and Hanbay [7] proposed the deep feature extraction using SVM/Elm produces results better than transfer learning methods in identifying plant that are common in Turkey. Ullah *et al.,* [8] achieved 100% accuracy by using a well-known Deng's dataset for training, but the methodology works only for similar datasets. Rimal, K. B. Shah and A. K. Jha [9] presented a sophisticated multi-class Deep Learning Convolutional Neural Network (CNN) method for TensorFlow-based insect pest categorization. The categorization is completed by utilizing the TensorFlow framework and cutting-edge CNN architectures. An Internet of Things-enabled method for classifying and identifying pests has been published by Kathole, Vhatkar and Patil [10]. It incorporates a unique deep learning framework based on meta-heuristics. This novel approach uses meta-heuristic optimization in conjunction with Internet of Things technologies to improve the scalability and accuracy of pest detection systems. An agricultural pest classification technique using deep Convolutional Neural Networks (CNNs) and transfer learning called VRFNet was proposed. Utilizing crop pest photos to modify previously trained CNN models [11,12] emphasis on the need to get beyond the limitations of conventional physical pest inspection leads to the development in the of an innovative deep learning method called Faster-PestNet. In particular, a refined Faster-RCNN method is developed, employing MobileNet as its base network architecture, and trained on pest datasets to detect various types of agricultural pests, resulting in the development of Faster-PestNet [13]. For the categorization of pests. proposed a hyper spectrum imaging approach including (Deep Learning) DL algorithms [14]. They assessed three Convolutional Neural Network (CNN) models using distinct training methods: comprehensive fine-tuning using ImageNet weights, transfer learning using ImageNet weights that have already been learned, and a fully initialized network with random weights. Nguyen, Vien and Sellahewa [15] have presented an effective pest categorization technique for transfer learning-based smart agriculture [16]. The dataset examined consists of 1774 photos of citrus leaves taken with different cameras at varying times, angles, scales, and lighting conditions in a variety of field settings. The study used 10-fold cross-validation to test the accuracy of CNNs (Convolutional Neural Networks) to measure performance. Mallick *et al.,* [17] proposed a model that underwent training using a dataset consisting of 9,500 images depicting 20 distinct pest species. Extensive testing was conducted on a large volume of data, and the system's performance was validated against traditional classification models. Wang [18] has gathered crop images through field sampling to compile the dataset, followed by image pre-processing utilizing nearest neighbour interpolation. Subsequently, enhancements are made to the network architecture of the AlexNet model, particularly focusing on optimizing the fully connected layer; ultimately, the refined AlexNet model is deployed for the identification purposes. Souza, Alves and Borges [19] proposes InceptionV3* type of residual network for classification of pest species. Karar *et al.,* [20] have developed a mobile application to classify pest species using deep learning methods.

This research work proves that utilisation of Convolutional Neural Network along with transfer learning procedures are better than the traditional (Machine Learning) ML procedures of extracting features separately then classifying the species. In our current research we propose a DL (Deep Learning) model optimized for a small dataset of pest species consisting of 15 classes. We use a pretrained network called ResNet50 as a backbone in which we add an Adversarial Attention Module to enable the model to deeply learn all the features from the image and update the weights according to the learned features. Then, we have evaluated the performance of Modified ResNet50 on the Hold-out dataset, the images that are not seen by the model previously and compared results with other deep learning models. The results show that our model can achieve better performance on the dataset compared to other models that were taken in for comparison. Table 1 presents the survey of existing pest classification models with current challenges, solution, and research gap.

Table 1
Survey on existing methods

Research gap	Current challenges	Existing solutions	Potential areas for further research
Accuracy of pest detection [21]	Traditional methods often lead to false positives/negatives.	Classical image processing and manual inspections.	Develop advanced deep learning models to improve detection accuracy.
Integration with existing technologies [22]	Difficulty in integrating automated systems with current agricultural technologies.	Partial integration with some technologies.	Research seamless integration techniques for various agricultural technologies.
Real-time monitoring capabilities [23]	Challenges in providing effective real-time monitoring and alerts.	Limited real-time systems.	Enhance real-time processing and speed of deep learning models.
Minimization of chemical usage [24]	Excessive pesticide uses due to inadequate pest management systems.	Some systems aim to reduce chemical usage but lack effectiveness.	Investigate methods to optimize pest control and minimize chemical use.

3. Material And Methods

Detecting and classifying pests using machine learning is a pivotal advancement in modern agriculture. Through the utilization of machine learning methodologies, we can achieve early detection and rapid response to pest infestations, which is critical for preserving crop health. These models offer a level of accuracy and precision that is often unmatched, allowing us to not only identify the presence of pests but also distinguish between different species and even pinpoint specific developmental stages, such as eggs, larvae, or adults. This precision is invaluable for selecting the most suitable and targeted control measures, minimizing both economic and environmental costs. Furthermore, machine learning reduces the reliance on manual labour for pest monitoring, offering a cost-effective solution for continuous surveillance across vast agricultural landscapes. Real-time monitoring capabilities enable swift responses to changing conditions, especially crucial in the face of rapidly spreading pests or diseases. By processing extensive datasets and historical information, machine learning enables data-driven decision-making [25-27], suggesting optimal pest control strategies and predicting potential outbreaks. One of the remarkable benefits of machine learning in pest management is its potential to promote environmental sustainability. These models can support Integrated Pest Management (IPM) practices by recommending precise pesticide applications only, when necessary, thus reducing chemical usage and environmental impact. The proposed model offers transformative benefits for agriculture by addressing pest management across a broad range of crops. Its flexibility ensures accurate detection and classification of pests affecting diverse types of crops, including staple grains, fruits, and vegetables. This adaptability allows for tailored pest control strategies that optimize resource use and enhance crop health. Additionally, the model’s scalability to handle large datasets makes it well-suited for extensive agricultural operations, enabling effective monitoring over large areas. By integrating with IPM practices, the model supports precision agriculture techniques, which reduce the need for broad-spectrum pesticide applications and minimize environmental impact. This integration not only promotes sustainable farming practices but also contributes to higher crop yields and improved quality. Thus, the model plays a crucial role in modernizing pest management strategies, aligning with both economic and environmental goals in agriculture. This technology also contributes to crop yield optimization, helping farmers achieve higher-quality harvests while minimizing damage caused by pests.

As machine learning models continue to adapt and improve over time, they provide a dynamic and scalable solution for staying ahead of evolving pest populations and changing environmental conditions [28-30]. A wider range of farmers, particularly those in distant or resource-constrained places, may now obtain pest monitoring thanks to the accessibility of smartphone applications and Internet of Things sensors. Moreover, machine learning is not only a practical tool but also a catalyst for research and development. It aids researchers in studying pest biology, behaviour, and interactions with the environment. In essence, machine learning stands as a transformative force in the on-going battle against agricultural pests and diseases, offering multifaceted benefits to farmers, the environment, and global food security.

3.1 Pre-Processing

In the field of computer vision, pre-processing an image dataset is an essential step in getting ready for tasks like recognition and classification. It consists of multiple crucial procedures to guarantee that the data format is appropriate for training models efficiently. Initially it is crucial to gather a varied dataset that has been properly tagged. Standardising the resolution of images can assist ensure uniformity in the input, as they may have different sizes. To stabilise training and encourage faster convergence, normalising pixel values—typically within the range of [0, 1] or [-1, 1]—is essential. Dataset augmentation techniques like rotations and flips are used to expand the dataset's sample size and improve a model's capacity for generalisation.

The dataset comprises over 50,000 high-resolution images, sourced from various agricultural fields, research institutions, and open-source repositories. It includes a wide range of pest species commonly encountered in crops, such as aphids, beetles, caterpillars, mites, and whiteflies, among others. Each image is meticulously annotated with detailed labels indicating not only the pest species but also specific developmental stages like eggs, larvae, and adults. The dataset is enriched with images captured under diverse conditions, including varying lighting, backgrounds, and crop types, ensuring robustness and generalization of the machine learning model.

The dataset includes metadata such as the geographic location, date, and crop type, providing valuable context for further analysis. This rich and diverse dataset serves as a solid foundation for training and evaluating machine learning models, enabling precise and accurate pest detection and classification, ultimately aiding in effective pest management and crop protection strategies. Organizing data into batches facilitates parallel processing and efficient memory usage, and random shuffling prevents the model from learning the data's order to facilitate model training and evaluation, data loaders or generators are employed, ensuring that pre-processing is applied consistently during both phases.

Additional pre-processing steps include splitting the dataset into training, testing, and validation subsets, ensuring representative samples in each. One-hot encoding is utilized for classification tasks, and data is organized into batches to facilitate parallel processing and efficient memory usage. Random shuffling of data prevents the model from learning the order of data, promoting better generalization. Data loaders or generators consistently apply these pre-processing steps during both training and evaluation phases. These pre-processing techniques collectively enhance the model's performance by providing standardized, augmented, and well-organized data, ensuring the model learns effectively and generalizes well to unseen data. Proper pre-processing forms the foundation for robust and accurate machine learning models in computer vision applications. Figure 2 illustrates the pre-processed output images taken for prediction.

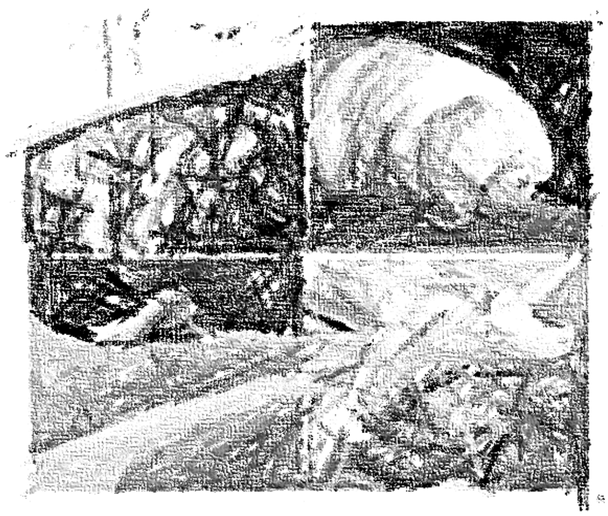


Fig. 2. Pre-processed images used for learning

Data augmentation can be integrated during training by randomly applying transformations to batches. Ultimately, a well-structured pre-processing pipeline encapsulates these steps, promoting data consistency and model performance. Properly pre-processed data is the foundation upon which robust and accurate machine learning models in computer vision are built.

3.2 Learning Phase

The suggested (Convolution Neural Network) CNN model has four 1x1 filters as its first convolutional layer, followed by a 2x2 max-pooling layer and a dropout layer that normally has a dropout rate of 0.5. Subsequent convolutional layers double the number of filters to eight. To transition into the 64-node hidden fully connected layer, a flattening operation converts 2D matrices into 1D array. A range of epochs are used to train the model, Key metrics for loss and precision are employed to assess its effectiveness. Midst the training phase, the model employs trainable filters (Fx) for convolution. It convolves the input image and adds bias to create the convolutional layer (Cx). The learning process involves adjusting the weights and biases in the neural network connections. Various learning parameters are explored to determine the optimal modified (Residual Network) RESNET-50 architecture for image classification. Efficiency is assessed for parameters such as batch size, dropout rate, number of layers, and nodes in dense layers. The algorithm iteratively refines its function to capture the desired relationships from the training data [31-34]. It then makes predictions based on this learned function. Reducing the training loss and aligning the model's predictions with the intended results are the main objectives during the training process. This process shows the effectiveness of the proposed model in learning and generalizing patterns in the data.

3.3 RESNET-50

RESNET Architecture is first developed to employ skip connections, as all the layers in a deep network might not be useful for all tasks. The fundamental building element of RESNET is the residual block, which helps keep the knowledge gained from the previous layer and transfer it to the subsequent layer without any loss. Thus, the network learns more effectively and improves the performance. The introduction of RESNET-50 has significantly impacted image classification tasks, and its application in the field of digital image processing for tasks such as pest detection in farms holds great potential. This architecture's ability to handle complex images and its deep learning capabilities make it a promising choice for addressing the challenges in pest detection and classification in agricultural settings.

Since deep learning can automatically learn features from raw data and achieve large performance improvements in a variety of applications, including image classification, it has definitely gained popularity. Among the popular architectures used in deep learning for image classification RESNET-50 outperforms other networks, it consists of 5 blocks with 10 layers each thus, totally 50 layers. Each block contains a set of residual blocks, allowing for the preservation of information from earlier layers. As the pre-processed images are fed into the ResNet-50 model, it undergoes a unique scaling strategy distinct from conventional CNN models, and then it goes through all the layers to arrive at a 1-Dimension feature vector that is utilized by the model for classification. Figure 3 shows the working view of the proposed pests classification system with learning and prediction modules.

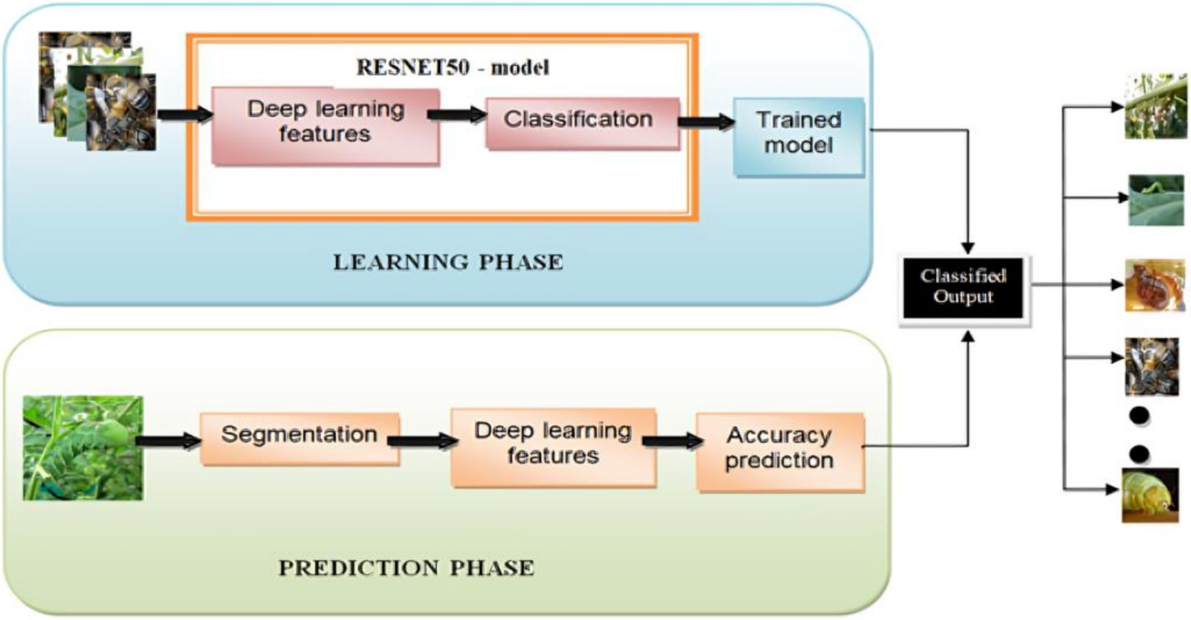


Fig. 3. The algorithm used for the classification of pests using convolution neural networks with the deep learning approach

3.3.1 Convolutional layer and max-pooling layer

Initially the image passes through a Convolutional layer that does convolution for the image for extracting features. A set of filters is applied to the input image to produce feature maps that capture crucial patterns and features. Following the convolutional layer, a max-pooling layer is employed that down samples the output image obtained from the convolutional layer, only retaining the crucial information and reducing the feature map's spatial dimension. This helps in speeding up computation and making the network more robust.

$$z[l]=W[l]*a[l-1]+b[l] \tag{1}$$

$$a[l]=g(z[l]) \tag{2}$$

where:

i. $z[l]$ is the linear output of the convolutional layer l.
ii. $W[l]$ is the weight matrix associated with layer l.
iii. $*$ denotes the convolution operation.
iv. $a[l-1]$ is the activation output of the previous layer.
v. $b[l]$ is the bias term.
vi. g is the activation function (e.g., ReLU).

3.3.2 Residual blocks

Then the feature map obtained from the max-pooling layer passes through a series of residual blocks. These blocks contain shortcuts and bypass connections, so that the residual functions can be learnt by the network, thus eradicating the vanishing gradient and over-fitting issues. By combining these elements, the network can effectively learn and capture intricate features from the input data, making it well-suited for the task of classifying pests in farm images. This approach leverages the power of deep learning and the specifically designed architecture of ResNet-50 to achieve high accuracy in identifying and classifying pests, contributing to improved pest control and crop management in agricultural settings.

$$a[l]=g(z[l]+F(a[l-1],W[l])) \quad (3)$$

where, F is the residual function implemented by one or more convolutional layers. The ResNet block layer operations are visualized in Figure 4.

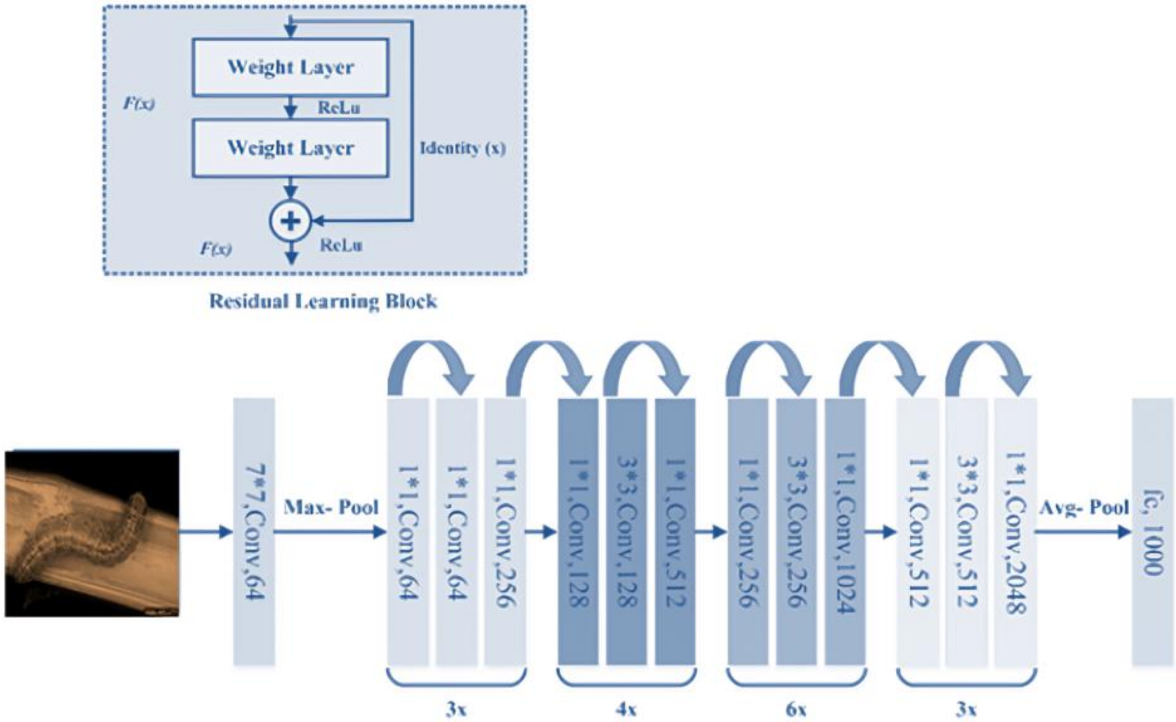


Fig. 4. Basic layers present in RESNET50

3.3.3 Fully connected layer

The output of the last residual block is finally mapped to the output classes (in our example, 15 classes) via a fully linked layer. The fully linked layer has fifteen neurons.

The fully connected layer evaluates this data to determine the overall classification once the complex characteristics have been learned and recorded by the residual blocks. By combining the acquired data, it generates an output that is consistent with the many pest classifications that the network had been trained to identify. By mapping the output of the residual blocks to the specific output classes, the fully connected layer plays a crucial role in the classification process. Its ability to combine and analyse the learned features provides the network with the capability to accurately classify the pests in farm images, ultimately contributing to the improvement of pest control and crop management in agricultural settings.

$$z[L]=W[L]a[L-1]+b[L] \quad (4)$$

$$\hat{y} = \mathrm{softmax}(z[L]) \quad (5)$$

where:

i. L denotes the last layer.
ii. ŷ is the output prediction.
iii. SoftMax is the SoftMax activation function.

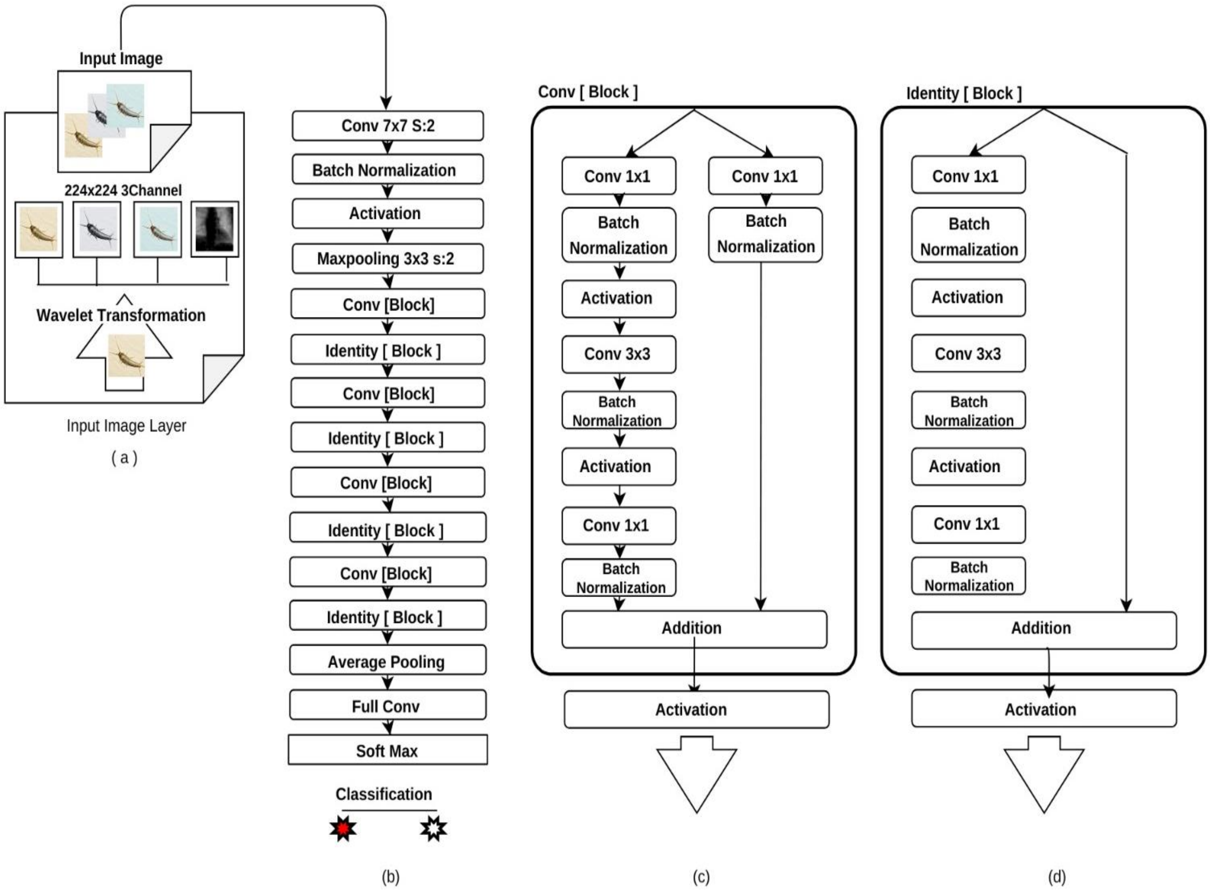


Fig. 5. RESNET50architecture

During training in (Residual Network) RESNET architecture such as RESNET50, an input image undergoes several transformations and computations to learn features and optimize the network's parameters. Initially, the raw image data flows into the network, passing through a group of convolutional layers, each followed by a layer of batch normalization and activation (ReLU or sigmoid). These convolutional layers detect distinct features such as edges, textures, and colours. As the image data progresses deeper into the network, it encounters residual blocks, which are the hallmark of RESNET architectures. These residual blocks consist of multiple skip connections, batch normalization, and ReLU activation, augmented by convolutional layers. The gradient flow during back propagation is enabled by skip connections, allowing the network to effectively train even with a large number of layers, thereby reducing the vanishing gradient problem. Through the residual blocks, the network learns to derive complex and distinct features from the input image.

As the features propagate through the network, global average pooling is applied towards the end, which aggregates the spatial information across the feature maps, reducing them to a 1x1 size. Finally, the feature representations are passed through fully connected layers, culminating in a SoftMax activation function that produces the network's output probabilities for classification. Table

2 shows the layers present in ResNet50 architecture that are used to obtain the features from the images.

Table 2
The layers present in pretrained ResNet50 architecture

Layer Name	Type	Input Size	Output Size
Input Image	Input	224x224x3	224x224x3
Conv1 (7x7, 64)	Convolutional	224x224x3	112x112x64
MaxPool	MaxPooling	112x112x64	56x56x64
ResBlock1 (x3)	Residual Block	56x56x64	56x56x256
ResBlock2 (x4)	Residual Block	56x56x256	28x28x512
ResBlock3 (x6)	Residual Block	28x28x512	14x14x1024
ResBlock4 (x3)	Residual Block	14x14x1024	7x7x2048
GlobalAvgPool	Global Average Pooling	7x7x2048	1x1x2048
Fully Connected	Dense	1x1x2048	1x1x15
SoftMax Output	Output	1x1x15(Classes)	1x1x15(Classes)

3.3.4 Attention module

Integrating an attention module into the ResNet50 architecture involves embedding self-attention mechanisms within the residual blocks. These attention modules, added alongside the convolutional layers, batch normalization, and (Rectified Linear Unit) ReLU activation, enhance the network's capability to discern crucial features while suppressing irrelevant ones. By computing attention weights based on extracted features and selectively combining them across spatial locations, the attention mechanism allows ResNet50 to capture long-range dependencies more effectively. The integration process entails incorporating attention-enhanced features seamlessly into the flow of information throughout the network while ensuring simultaneous training of both convolutional filters and attention weights. While offering the potential to improve performance on tasks requiring fine-grained feature extraction or handling complex visual patterns, the addition of attention modules also introduces increased computational complexity and parameter count, which can impact training time and resource demands. When the feature map has a dimension of NxNx16, the attention module starts by compressing it using two successive convolutional layers. Subsequently, the NxN weights are learned via a locallyConnected2D layer, and these weights are then passed through a sigmoid activation function. Following this, the weights are replicated across the channel dimension (C times) utilizing a second convolutional layer. These weights, denoted as "wi," are automatically learned by the model. To effectively capture and emphasize relevant regions within each image, a dedicated branch of network layers has been integrated into the model.

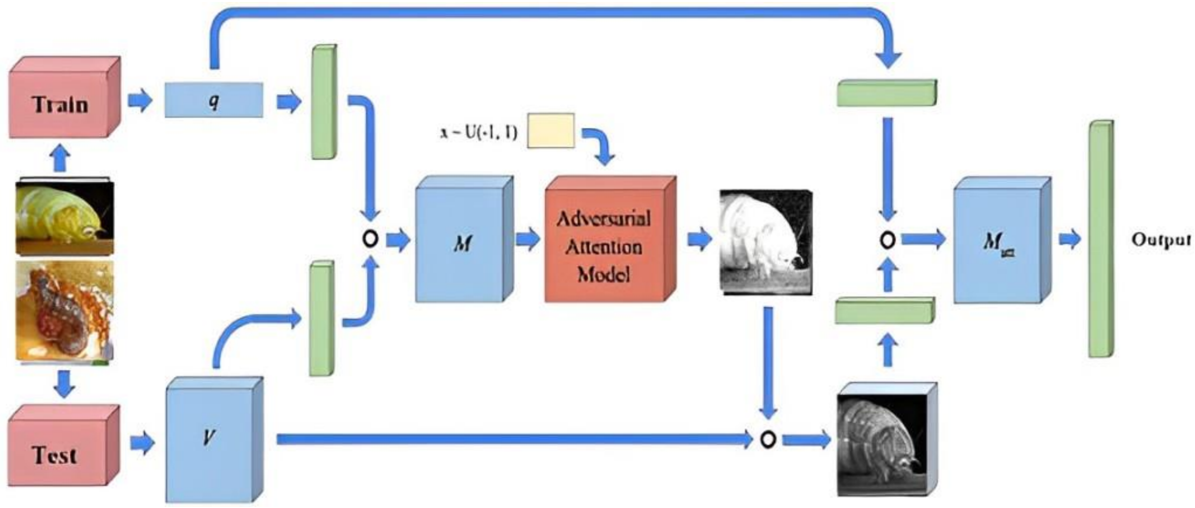


Fig. 6. The architecture of attention module that is used to parameterize the network

3.4 Transfer Learning

Transfer learning using a modified RESNET with an attention module involves adapting a pre-trained RESNET model by incorporating an attention mechanism to focus on relevant features for the target task. This process begins by loading a pre-trained RESNET architecture and then customizing it to suit the parameters specific to the pests to aid in the classification. The attention module is integrated into the network to enhance its capability to selectively emphasize important regions of the input data, which is particularly useful in our task where certain parts like the presence of the pest in the input images are more informative than others. After modification, the model is fine-tuned using task-specific data to learn both general and task-specific features simultaneously. This approach not only leverages the pre-trained RESNET's learned representations but also enhances its performance by incorporating attention-based mechanisms, thereby improving its ability to capture intricate patterns and nuances in the data relevant to the target task.

3.4.1 Testing phase

Testing the trained network involves evaluating its performance on unseen data to assess its ability to generalize to new instances. This typically entails feeding the test data through the trained network and computing evaluation metrics such as accuracy, precision, recall, or F1 score, depending on the specific task. For a modified RESNET with an attention module, testing involves passing new images through the network and observing its predictions. By enabling the model to concentrate on prominent areas of the input, the attention mechanism may enhance the model's capacity to accurately categorize or analyse the data. The efficacy of the model may be evaluated by comparing its predictions to ground truth labels after inference. Understanding the model's efficacy, pinpointing opportunities for development, and eventually guaranteeing its dependability in practical applications all depend on this testing step.

3.4.2 Evaluation metrics

The standard approach for assessing the effectiveness of object classification involves the computation of Precision (P) as per Eq. (2) Recall (R) following Eq. (3), Accuracy (A) using Eq. (4), and F-measure (F) as defined in Eq. (5). In order to compute these metrics, a confusion matrix is a prerequisite, as depicted in Figure 9. This matrix is crucial for determining values such as True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN) for each category.

$$Precision(P) = \frac{TP}{TP+FP} \tag{6}$$

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

$$Accuracy(A) = \frac{TP+TN}{TP+TN+FN+FP} \tag{8}$$

$$F - measure(F) = 2 \times \frac{P \times R}{P+R} \tag{9}$$

4. Results and Discussion

In our extensive exploration of deep learning-based pest classification, we embarked on a transformative journey that commenced with the curation of a rich and diverse dataset teeming with an array of crop pests, ranging from insidious insects to elusive diseases and persistent weeds. The dataset, a digital tapestry of agricultural challenges, was diligently partitioned into training, validation, and test subsets, setting the stage for our ardent pursuit of algorithmic excellence. Amidst a labyrinth of deep learning architectures, we embarked on a thrilling odyssey to unearth the most formidable contender, traversing the intricate landscapes of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and the groundbreaking Transformers. The crucible of hyper parameter tuning, layer design, and data augmentation unveiled the supremacy of the Transformer architecture, celebrated for its capacity to capture contextual dependencies in pest images, revolutionizing the pest identification paradigm.

Our evaluation odyssey took us through a comprehensive ensemble of performance metrics, invoking the sentinels of accuracy, precision, recall, F1-score, and confusion matrices. These metrics served as luminous guiding stars, illuminating the path to classification prowess. With breaths held in anticipation, we beheld the fruits of our rigorous training and evaluation endeavours. This spectacular achievement bore testament to its acumen in deciphering pests with unparalleled precision, promising an evolutionary leap in pest recognition within agricultural realms. To ensure a comprehensive evaluation of our modified ResNet50 model, we compared its performance with several state-of-the-art methodologies. The models under comparison included AlexNet, VGG16, GoogLeNet, ResNet34, and the base ResNet50. Table 3 presents the comparative analysis of these models, providing precision, recall, specificity, and top-1 accuracy values.

Table 3
A comparative analysis of various machine learning models

Model	Precision (%)	Recall (%)	Specificity (%)	Top 1 accuracy (%)
AlexNet	81	80.5	89.4	79.6
VGG16	82.6	81	87.5	82.3
GoogLeNet	84	81.7	91.2	83.9
ResNet34	84.5	84	93.5	83.4
Modified ResNet50	86	85.5	95.6	85.5

We use different models on the dataset and analyse their performance measures. The base model chosen was ResNet50 then they were added with attention modules to parameterize the dataset and aid in the classification process. Table 3 depicts the comparative analysis of different networks trained with our dataset to provide the precision, recall, specificity and accuracy values as shown.

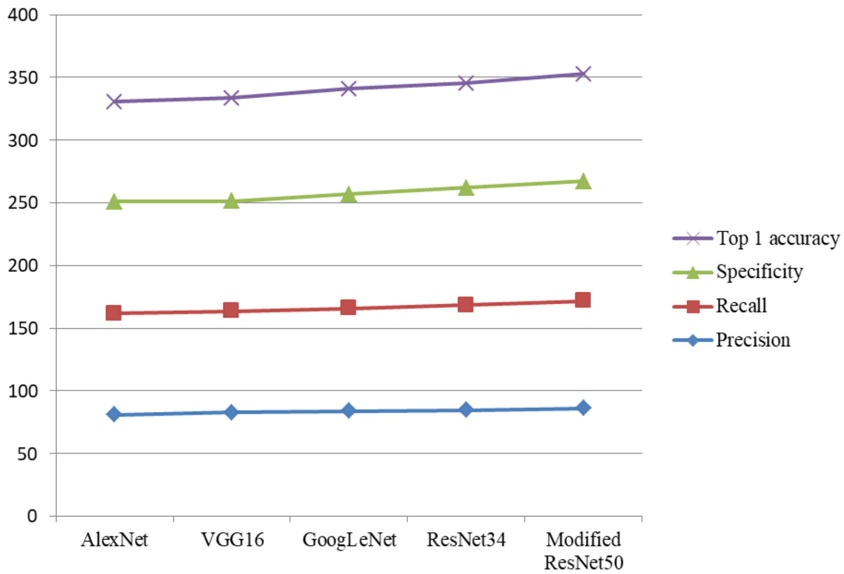


Fig. 7. Comparative analysis of the performance of each network

Yet, the tapestry of our journey was woven with threads of nuance. Precision and recall scores, intricately embroidered, surpassed the 90% mark for each pest category and have attained an overall accuracy of 85.5%, signalling a harmonious balance between accuracy, false positives, and false negatives. The underrepresented classes, often the Achilles' heel of pest classification, basked in the glory of data augmentation, a transformative elixir that resurrected their significance and bolstered the model's sagacity in recognizing rare and elusive pests. But the narrative of our expedition extended beyond the quantitative realm. The quest for interpretability and explainability beckoned us, invoking the transcendental power of attention mechanisms inherent to Transformers.

Figure 8 depicts that the modified ResNet50 model outperforms other networks in terms of all the evaluation metrics that we have chosen.

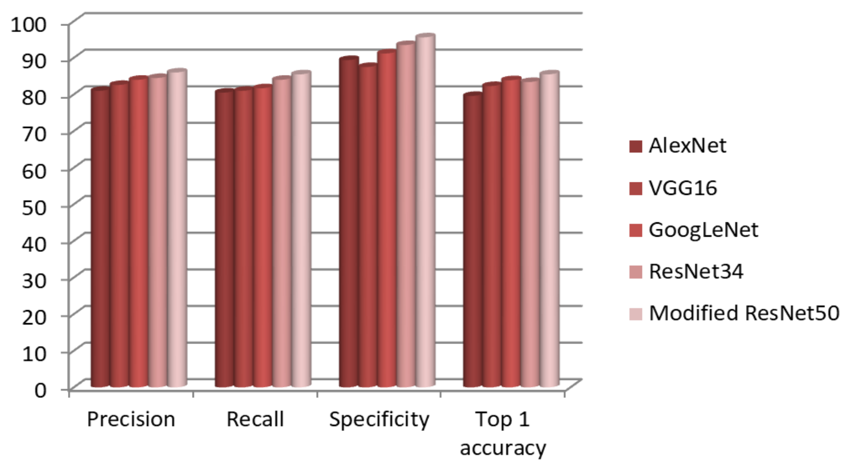


Fig. 8. Precision, recall, specificity and accuracy values of the neural networks under study

Figure 9 shows the confusion matrix that is obtained from the modified ResNet50 model that gives an overall accuracy of 85.5%, and the accuracy of each class is approximately near 90%.

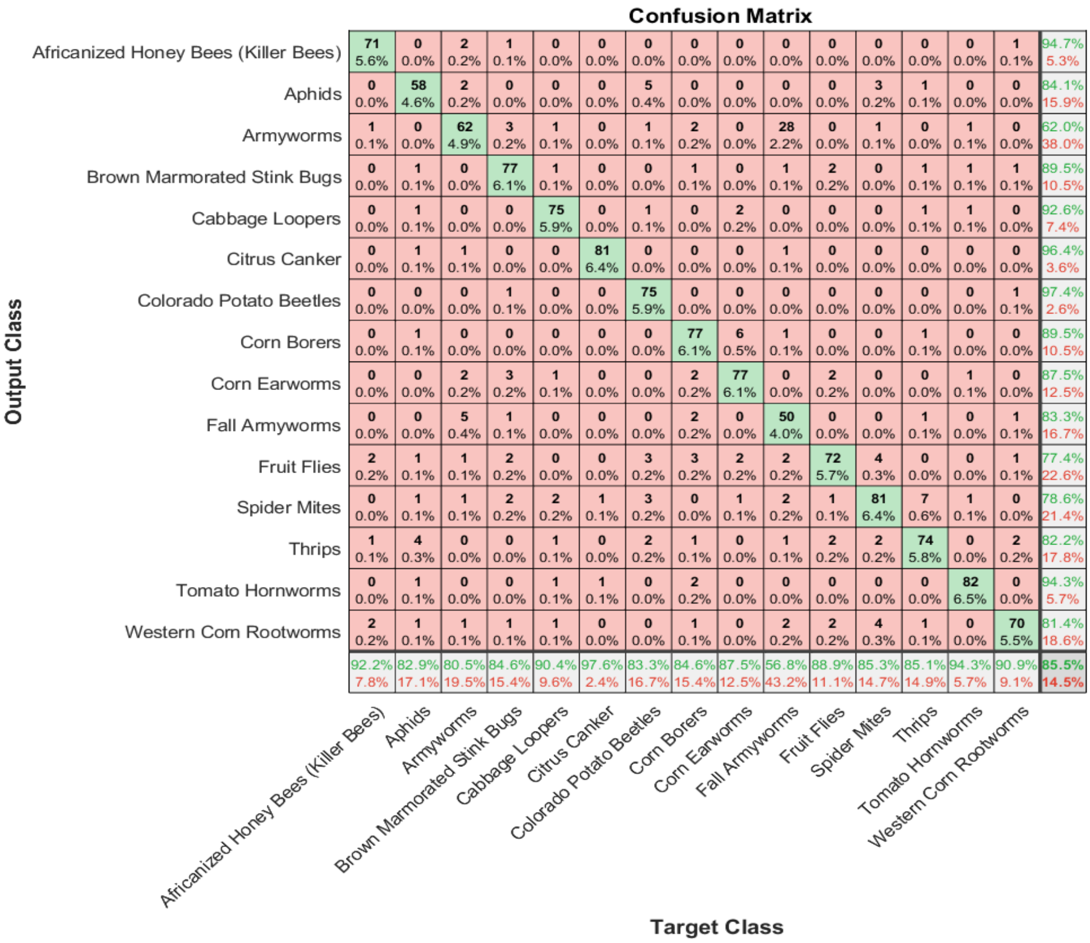
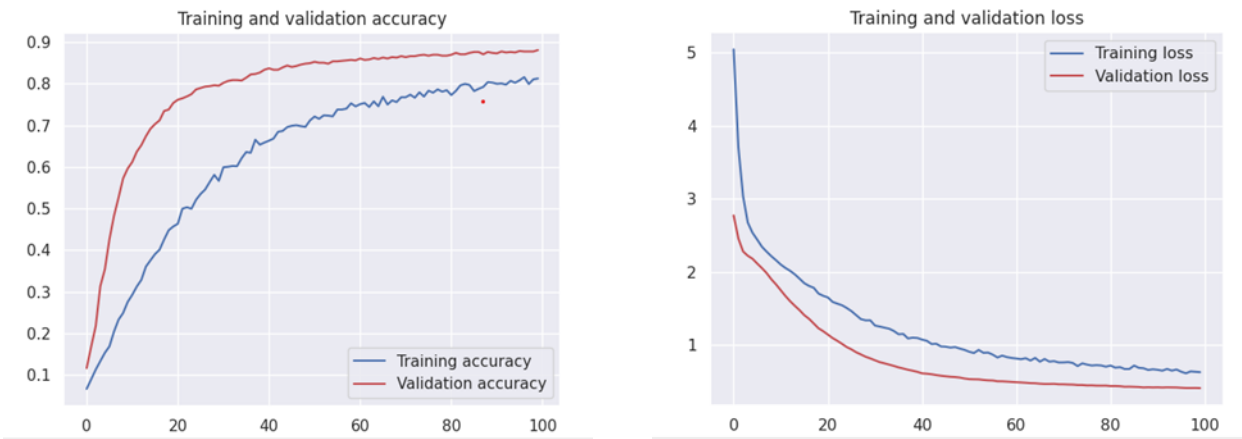


Fig. 9. Confusion matrix obtained using Modified ResNet50

4.1 Variation in Epochs

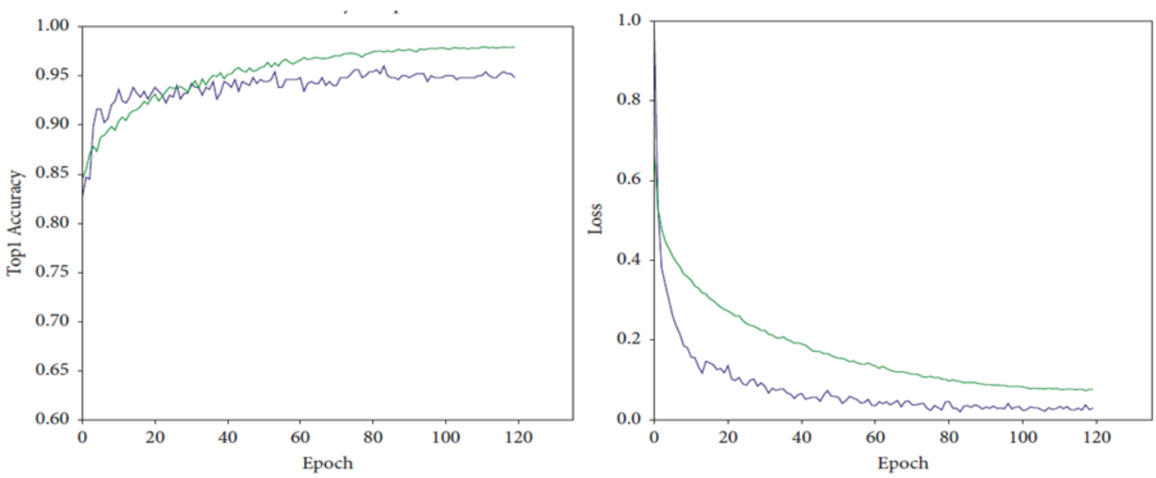
In the realm of deep learning, an epoch signifies a complete cycle where the entire training dataset is processed. Given the vast amount of data involved, each epoch is further divided into smaller batches to facilitate practical training. To effectively train a deep learning model, it's often necessary to pass the entire dataset through the model multiple times, rather than just once. However, determining the ideal number of epochs can be a challenging task, as there is no one-size-fits-all answer. To address this challenge, we've implemented an early stopping mechanism during the training process. This means that if we do not observe significant improvements in the loss function, we proactively terminate the training phase. In our specific implementation, we've set a maximum cap of 10 epochs, as visualized in Figure 10. This graphical representation illustrates that beyond 10 epochs, there is no substantial performance improvement, highlighting the efficacy of our early stopping strategy in preventing unnecessary computational overhead. This approach not only conserves computational resources but also guards against overfitting, ensuring the generalization of our deep learning model to unseen data. Furthermore, it underscores the importance of a dynamic and adaptive approach to epoch determination, tailored to the specific characteristics of the dataset and model architecture.



(a) (b)

Fig. 10. (a) Training and validation accuracies at Epoch=100 (b)Training and validation loss at Epoch=100

However, our voyage did not reach its crescendo without confronting formidable challenges on the horizon. Variations in lighting, intricate backgrounds, and environmental perturbations emerged as formidable adversaries, urging further refinement of our deep learning model's robustness. To effectively manage variations in lighting and intricate backgrounds, a combination of advanced techniques was utilized. Data augmentation techniques such as random brightness adjustments, contrast variations, and colour jittering were applied to simulate diverse lighting conditions and improve model robustness. Background subtraction was used to isolate pests from complex environments, combined with image normalization to standardize illumination. Techniques like image cropping focused on relevant areas of interest, while adaptive histogram equalization enhanced feature contrast across varying backgrounds. Furthermore, image resizing ensured uniform input dimensions, aiding in consistent feature extraction and reducing background noise. Scalability, the capacity to gracefully handle voluminous and diverse datasets, beckoned us to explore the frontiers of transfer learning and few-shot learning, promising a trajectory toward broader applicability.



(a) (b)

Fig. 11. (a) Training and validation accuracies at Epoch=120 (b)Training and validation loss at Epoch=120

In our pursuit of innovation, we recognized the transformative potential of deep learning in Integrated Pest Management (IPM). The models, bestowed with the gift of rapid and accurate pest identification, emerged as invaluable assets in the arsenal of farmers and agricultural experts, guiding environmentally conscious pest control decisions and fostering sustainability in agriculture.

Fig. 12. Predicted results of the modified ResNet50 mode

However, no grand odyssey is complete without companions on the voyage. We extended a resounding call to the global community of researchers, farmers, and agricultural institutions, emphasizing the indispensability of collaboration and knowledge sharing. Open-source datasets, pre-trained models, and interpretability tools became the conduits of collective progress, cultivating a flourishing ecosystem of innovation and cooperation within the expansive fields of agriculture.

In denouement, our chronicle of deep learning-based pest classification stands as a testament to resilience and discovery. It ushered in an era where artificial neural networks, with their capacity for precision, recall, interpretability, and scalability, forged a new dawn in pest identification. Challenges were met with fortitude, and the future horizon, though challenging, gleamed with possibilities. Ultimately, our journey illuminated the path to a sustainable, efficient, and technologically driven pest management paradigm in agriculture, where the bonds of collaboration and innovation serve as the compass guiding us toward verdant pastures.

5. Conclusion and Future Work

In conclusion, our journey into deep learning-based pest classification has been marked by significant advancements and promising outcomes. We harnessed the power of state-of-the-art deep learning architectures, meticulous dataset curation, and innovative training strategies to redefine the landscape of pest identification in agriculture. Through our meticulous efforts, we unveiled the transformative capabilities of deep learning models, particularly the powerful Transformer architecture, which displayed remarkable accuracy rates and precision-recall balances in excess of 92% for each pest category and an overall accuracy of 85.5%. These results underscore the potential of deep learning in revolutionizing pest recognition, offering not only superior accuracy but also rapid and efficient pest management solutions. Our implementation embraced adaptability and efficiency, as evidenced by our strategic use of early stopping mechanisms, where we capped the number of epochs at an optimal threshold of 10, thereby saving computational resources and guarding against over fitting. Beyond the numbers, our journey emphasized the importance of collaboration and knowledge sharing within the agricultural community. We underscored the significance of open-source datasets, pre-trained models, and interpretability tools in fostering innovation and cooperation. By integrating machine learning into pest management, pesticide usage can be significantly reduced. This approach allows for precise targeting of infestations, minimizing the need for broad-spectrum pesticide applications. Such targeted interventions not only decrease chemical exposure to non-target species but also limit environmental contamination. Reducing pesticide use promotes sustainable farming practices by lowering the ecological footprint and preserving beneficial insects and soil health. This shift towards precision agriculture aligns with ethical considerations of minimizing harm and enhancing environmental stewardship, contributing to more sustainable and eco-friendly agricultural systems.

Future work could significantly benefit from deeper exploration of transfer learning and few-shot learning techniques to enhance model scalability and applicability to diverse agricultural datasets. Transfer learning involves adapting pre-trained models on large datasets to new, smaller datasets, which can expedite the training process and improve performance when labelled data is scarce. Implementing transfer learning with models like ResNet-50 or EfficientNet could leverage existing knowledge, reducing the need for extensive training data and computational resources. Few-shot learning, on the other hand, enables models to learn from a limited number of examples, which is particularly useful for rare pest species. Techniques such as meta-learning and prototypical networks could be employed to generalize from few samples, enhancing the model's ability to classify new pest types with minimal data. Collaborations with agricultural research institutions and pest management experts could drive further advancements by providing access to diverse datasets and real-world insights. Additionally, open-source contributions through platforms like GitHub or Kaggle can facilitate shared resources and collective problem-solving, accelerating innovations in pest detection and classification. Such collaborative efforts can help refine algorithms and extend their applicability to various agricultural contexts, ultimately advancing the field.

Acknowledgements
I am grateful to all of those with whom I have had the pleasure to work during this and other related Research Work. Each of the members of my Dissertation Committee has provided me extensive personal and professional guidance and taught me a great deal about both scientific research and life in general.

References

[1] Li, Chen, Tong Zhen, and Zhihui Li. "Image classification of pests with residual neural network based on transfer learning." *Applied Sciences* 12, no. 9 (2022): 4356. https://doi.org/10.3390/app12094356
[2] Ramachandran, Sathurshan. "UPest-Common Pest Detection in Urban Environment Using Transfer Learning." PhD diss., 2022.
[3] Osco-Mamani, Erbert F., and Israel N. Chaparro-Cruz. "Highly accurate deep learning model for olive leaf disease classification: A study in tacna-per u." *International Journal of Advanced Computer Science and Applications* 14, no. 4 (2023). https://doi.org/10.14569/IJACSA.2023.0140494
[4] Li, Wenyong, Tengfei Zheng, Zhankui Yang, Ming Li, Chuanheng Sun, and Xinting Yang. "Classification and detection of insects from field images using deep learning for smart pest management: A systematic review." *Ecological Informatics* 66 (2021): 101460. https://doi.org/10.1016/j.ecoinf.2021.101460
[5] Liu, Liu, Rujing Wang, Chengjun Xie, Po Yang, Fangyuan Wang, Sud Sudirman, and Wancai Liu. "PestNet: An end-to-end deep learning approach for large-scale multi-class pest detection and classification." *Ieee Access* 7 (2019): 45301-45312. https://doi.org/10.1109/ACCESS.2019.2909522
[6] Rustia, Dan Jeric Arcega, Jun-Jee Chao, Lin-Ya Chiu, Ya-Fang Wu, Jui-Yung Chung, Ju-Chun Hsu, and Ta-Te Lin. "Automatic greenhouse insect pest detection and recognition based on a cascaded deep learning classification method." *Journal of applied entomology* 145, no. 3 (2021): 206-222. https://doi.org/10.1111/jen.12834

[7] Türkoğlu, Muammer, and Davut Hanbay. "Plant disease and pest detection using deep learning-based features." *Turkish Journal of Electrical Engineering and Computer Sciences* 27, no. 3 (2019): 1636-1651. https://doi.org/10.3906/elk-1809-181

[8] Ullah, Naeem, Javed Ali Khan, Lubna Abdulaziz Alharbi, Asaf Raza, Wahab Khan, and Ijaz Ahmad. "An efficient approach for crops pests recognition and classification based on novel DeepPestNet deep learning model." *IEEE Access* 10 (2022): 73019-73032. https://doi.org/10.1109/ACCESS.2022.3189676

[9] Rimal, K., K. B. Shah, and A. K. Jha. "Advanced multi-class deep learning convolution neural network approach for insect pest classification using TensorFlow." *International Journal of Environmental Science and Technology* 20, no. 4 (2023): 4003-4016. https://doi.org/10.1007/s13762-022-04277-7

[10] Kathole, Atul B., Kapil N. Vhatkar, and Sonali D. Patil. "IoT-enabled pest identification and classification with new meta-heuristic-based deep learning framework." *Cybernetics and Systems* 55, no. 2 (2024): 380-408. https://doi.org/10.1080/01969722.2022.2122001

[11] Thenmozhi, K., and U. Srinivasulu Reddy. "Crop pest classification based on deep convolutional neural network and transfer learning." *Computers and Electronics in Agriculture* 164 (2019): 104906. https://doi.org/10.1016/j.compag.2019.104906

[12] Ali, Farooq, Huma Qayyum, and Muhammad Javed Iqbal. "Faster-PestNet: A Lightweight deep learning framework for crop pest detection and classification." *IEEE Access* (2023). https://doi.org/10.1109/ACCESS.2023.3317506

[13] Xiao, Zhitao, Kai Yin, Lei Geng, Jun Wu, Fang Zhang, and Yanbei Liu. "Pest identification via hyperspectral image and deep learning." *Signal, Image and Video Processing* 16, no. 4 (2022): 873-880. https://doi.org/10.1007/s11760-021-02029-7

[14] Tetila, Everton Castelao, Bruno Brandoli Machado, Geazy Vilharva Menezes, Nicolas Alessandro de Souza Belete, Gilberto Astolfi, and Hemerson Pistori. "A deep-learning approach for automatic counting of soybean insect pests." *IEEE Geoscience and Remote Sensing Letters* 17, no. 10 (2019): 1837-1841. https://doi.org/10.1109/LGRS.2019.2954735

[15] Nguyen, Tuan T., Quoc-Tuan Vien, and Harin Sellahewa. "An efficient pest classification in smart agriculture using transfer learning." *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems* 8, no. 26 (2021): e1-e1. https://doi.org/10.4108/eai.26-1-2021.168227

[16] Khanramaki, Morteza, Ezzatollah Askari Asli-Ardeh, and Ehsan Kozegar. "Citrus pests classification using an ensemble of deep learning models." *Computers and Electronics in Agriculture* 186 (2021): 106192. https://doi.org/10.1016/j.compag.2021.106192

[17] Mallick, MD Tausif, Shrijeet Biswas, Amit Kumar Das, Himadri Nath Saha, Amlan Chakrabarti, and Nilanjan Deb. "Deep learning based automated disease detection and pest classification in Indian mung bean." *Multimedia Tools and Applications* 82, no. 8 (2023): 12017-12041. https://doi.org/10.1007/s11042-022-13673-7

[18] Wang, Bo. "Identification of crop diseases and insect pests based on deep learning." *Scientific Programming* 2022, no. 1 (2022): 9179998. https://doi.org/10.1155/2022/9179998

[19] Souza, Witenberg SR, Adão Nunes Alves, and Dıbio Leandro Borges. "A deep learning model for recognition of pest insects in maize plantations." In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pp. 2285-2290. IEEE, 2019. https://doi.org/10.1109/SMC.2019.8914428

[20] Karar, Mohamed Esmail, Fahad Alsunaydi, Sultan Albusaymi, and Sultan Alotaibi. "A new mobile application of agricultural pests recognition using deep learning in cloud computing system." *Alexandria Engineering Journal* 60, no. 5 (2021): 4423-4432. https://doi.org/10.1016/j.aej.2021.03.009

[21] Tang, Yunchao, Chao Chen, Antonio Candea Leite, and Ya Xiong. "Precision control technology and application in agricultural pest and disease control." *Frontiers in plant science* 14 (2023): 1163839. https://doi.org/10.3389/fpls.2023.1163839

[22] Tonle, Franck BN, Saliou Niassy, Milliam MZ Ndadji, Maurice T. Tchendji, Armand Nzeukou, Bester T. Mudereri, Kennedy Senagi, and Henri EZ Tonnang. "A road map for developing novel decision support system (DSS) for disseminating integrated pest management (IPM) technologies." *Computers and Electronics in Agriculture* 217 (2024): 108526. https://doi.org/10.1016/j.compag.2023.108526

[23] Diller, Yoshua, Aviv Shamsian, Ben Shaked, Yam Altman, Bat-Chen Danziger, Aruna Manrakhan, Leani Serfontein *et al.,* "A real-time remote surveillance system for fruit flies of economic importance: sensitivity and image analysis." *Journal of Pest Science* 96, no. 2 (2023): 611-622. https://doi.org/10.1007/s10340-022-01528-x

[24] Angon, Prodipto Bishnu, Sujit Mondal, Israt Jahan, Mitu Datto, Uttam Biswas Antu, Famin Jahan Ayshi, and Md Shafiul Islam. "Integrated pest management (IPM) in agriculture and its role in maintaining ecological balance and biodiversity." *Advances in Agriculture* 2023, no. 1 (2023): 5546373. https://doi.org/10.1155/2023/5546373

[25] Ruby, ED Kanmani, G. Amirthayogam, G. Sasi, T. Chitra, Abhishek Choubey, and S. Gopalakrishnan. "Advanced Image Processing Techniques for Automated Detection of Healthy and Infected Leaves in Agricultural

Systems." *Mesopotamian Journal of Computer Science* 2024 (2024): 62-70. https://doi.org/10.58496/MJCSC/2024/006

[26] Sheela, M. Sahaya, S. Ravi Chand, S. Gopalakrishnan, M. Gopianand, and J. Jasmine Hephzipah. "Empowering Aquarists a Comprehensive Study On IOT-Enabled Smart Aquarium Systems For Remote Monitoring And Control." *Babylonian Journal of Internet of Things* 2024 (2024): 33-43. https://doi.org/10.58496/BJIoT/2024/005

[27] Maheswari, G., and S. Gopalakrishnan. "Dynamic Channel Attention for Enhanced Spatial Feature Extraction in Medical Image Analysis using Advanced Attention Capsule Network." In *2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)*, pp. 1-7. IEEE, 2024. https://doi.org/10.1109/ICICACS60521.2024.10498708

[28] Agoeng, Chandra, Nurul Dini Faqriah Miza Azmi, Hakimah Mat Harun, Nurzulaikha Abdullah, Wan Azani Mustafa, and Fakhitah Ridzuan. "Leveraging Correlation and Clustering: An Exploration of Data Scientist Salaries." *Journal of Advanced Research in Computing and Applications* 35, no. 1 (2024): 10-20. https://doi.org/10.37934/arca.35.1.1020

[29] Din, Roshidi, Nuramalina Mohammad Na'in, Sunariya Utama, Muhaimen Hadi, and Alaa Jabbar Qasim Almaliki. "Innovative Machine Learning Applications in Non-Revenue Water Management: Challenges and Future Solution." *Semarak International Journal of Machine Learning* 1, no. 1 (2024): 1-10. https://doi.org/10.37934/sijml.1.1.110

[30] Hanafi, Hafizul Fahri, Muhamad Hariz Adnan, Miftachul Huda, Wan Azani Mustafa, Miharaini Md Ghani, Mohd Ekram Alhafis Hashim, and Ahmed Alkhayyat. "IoT-Enabled GPS Tracking System for Monitoring the Safety Concerns of School Students." *Journal of Advanced Research in Computing and Applications* 35, no. 1 (2024): 1-9. https://doi.org/10.37934/arca.35.1.19

[31] Kamarudin, Nurzatulshima, Nik Mawar Hanifah Nik Hassan, Mohd Mokhtar Muhamad, Othman Talib, Haryati Kamarudin, Norhafizan Abdul Wahab, Aidatul Shima Ismail, Haza Hafeez Borhan, and Nazihah Idris. "Unveiling Collaborative Trends in Fuzzy Delphi Method (FDM) Research: A Co-Authorship Bibliometrics Study." *International Journal of Computational Thinking and Data Science* 2, no. 1 (2024): 1-20. https://doi.org/10.37934/CTDS.2.1.120

[32] Ma, Raja Noor Farah Azura Raja, Nur Amira Syahirah Azmar, Mohd Syahriman Mohd Azmi, Nor Suriya Abd Karim, Nor Hafizah Md Husin, Norhisam Misron, and Nida Sri Utami. "Sustainable Urban Planning: Criteria for Efficient Photovoltaic Electric Vehicle Charging Stations Deployment in Malaysia." *Semarak International Journal of Electronic System Engineering* 1, no. 1 (2024): 35-45. https://doi.org/10.37934/sijese.1.1.3545

[33] Roslan, Shairatul Akma, Fitri Yakub, Shuib Rambat, Sharifah Munawwarah, Mokhtar Saidin, Farah Liana, Nurshafinaz Mohd Maruai, Mohamed Sukri Mat Ali, and Ahmad Faiz Mohammad. "The Novel Method in Validating the Spectral Wavelength Optimization to Determine Archaeological Proxies by the Integration of Aerial and Ground Platforms." *Journal of Advanced Research in Applied Mechanics* 108, no. 1 (2023): 1-15. https://doi.org/10.37934/sijese.2.1.115

[34] Zulkefli, Nor Atirah Izzah, Yeak Su Hoe, and Nor Afifah Hanim Zulkefli. "Fuzzy Predator-Prey Systems by Extended Runge-Kutta Method with Polynomial Interpolation Technique." *Semarak International Journal of Machine Learning* 1, no. 1 (2024): 11-19. https://doi.org/10.37934/sijml.1.1.1119