Shahid Bahonar University of Kerman

Biomechanism and Bioenergy Research

Online ISSN: 2821-1855
Homepage: https://bbr.uk.ac.ir

Iranian Society of Agricultural Machinery Engineering and Mechanization

Evaluating the Integration of Deep Neural Networks with Machine Learning Classification Algorithms to Detect and Classify Eight Pest Species in Tomato and Wheat Crops

Mohammad ghorbani[1], Kazem jafarinaeimi[1]✉, Hossein Maghsoudi[1], Bita Negarestani[1], Mohammad Maharlooei[2]

[1] Biosystems Engineering Department, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran.

[2] University of California, Agriculture and Natural Resources, CA, United States

✉ Corresponding author: jafarinaeimi@uk.ac.ir

ARTICLE INFO

Article type:

Research Article

Article history:

Received 11 March 2026

Received in revised form 07 April 2026

Accepted 14 May 2026

Available Online 30 June 2026

Keywords:

Tomato pest, Machine learning, Deep learning, Feature extraction, Transfer learning, Grad-CAM

ABSTRACT

Recognizing the challenges associated with conventional pest identification methods, which are labor-intensive, detrimental, and restrictive, the development of effective predictive models for accurate crop pest identification is paramount. This research utilized machine learning and deep learning techniques to extract features and classify tomato pests. Transfer learning was applied to a dataset of 1304 images of eight distinct tomato pests to expedite training. Critical pest characteristics were extracted employing convolutional neural network models, including AlexNet, GoogLeNet, ResNet101, and VGG16. Naïve Bayesian, Support Vector Machine, and K-Nearest Neighbors. With data augmentation and image preprocessing, ResNet101 achieved the highest performance, attaining 92.4% accuracy and surpassing the other models. The classification capabilities of ResNet101 and the SVM classifier were evaluated employing F1 scores across eight pest species (Bemisia argentifolii, Helicoverpa armigera, Myzus persicae, Spodoptera exigua, Spodoptera litura, Thrips palmi, Tetranychus, and Zeugodacus cucurbitae), yielding F1 scores of 69/74%, 92/87%, 88/88%, 75/69%, 52/65%, 76/9%, 35/35%, 72%, and 100%, respectively. The integration of machine learning and deep learning methodologies enables faster pest detection for experts and farmers, improves feature extraction, shortens training time, and ultimately increases crop yields and reduces economic losses.

Cite this article: Ghorbani, M., Jafarinaeimi, K., Maghsoudi, H., Negarestani, B. & Maharlooei, M. (2026). Evaluating The Integration of Deep Neural Networks with Machine Learning Classification Algorithms to Detect and Classify Eight Pest Species in Tomato and Wheat Crops. *Biomechanism and Bioenergy Research*, *5*(2), 46-67. https://doi.org/10.22103/bbr.2026.27265.1151

© The Author(s).

Publisher: Shahid Bahonar University of Kerman

DOI: https://doi.org/10.22103/bbr.2026.27265.1151

INTRODUCTION

Tomatoes (Solanum lycopersicum) are the most produced vegetable in the world, with a production volume of 189 million tons in 2021. Fresh, processed, or tomato-based products are prevalent in human diets and cuisines across many countries, providing essential nutritional benefits. For example, tomatoes are the second most-consumed vegetable in the United States, with a per capita availability of 14.3 kilograms. Tomatoes exhibit a strong nutritional profile and considerable versatility, making them suitable for producing a variety of food products (Bergougnoux, 2014). Consequently, tomato cultivation is an essential component of the global food supply chain, helping meet the food and nutritional needs of a growing global population (Amirahmadi et al., 2023). However, if not promptly managed, tomato crops are susceptible to various pests and diseases, which can substantially diminish yields and fruit quality (Jin et al., 2024; Wakil et al., 2018).

Insect pests can infest tomato crops at all stages of growth, from early development to late maturity, adversely affecting leaves, flowers, and fruits. This infestation may result in crop damage, diminished yields, and substandard fruit quality, impacting overall production and farmers' income (Depenbusch et al., 2025) and compromising the nutritional quality of human food. Throughout the annals of agricultural history, farmers have consistently encountered the challenges posed by herbivorous and disease-transmitting insects, collectively called pests, thereby necessitating a diverse array of control strategies (Oerke, 2006).

Conventional pest detection and classification methods frequently rely on manual inspection, which is time-consuming, labor-intensive, and error-prone. Furthermore, precise identification of insect pest species requires a high level of expertise or domain knowledge, and inaccuracies in identification can result in significant waste of resources, effort, and financial investment (Praharsha et al., 2024). While conventional agricultural practices use pesticides for pest management, this approach causes environmental pollution and poses hazards to consumer health. Agricultural specialists generally perform disease and pest identification; nonetheless, this process is often lengthy and expensive. Pest control strategies vary by species classification, and many species can carry or transmit plant viruses. Consequently, detecting various pests in extensive agricultural fields often poses challenges.

In recent years, advancements in computer vision and deep learning (DL) have facilitated the creation of automated pest recognition systems employing image analysis techniques (Altuntaş & Kocamaz, 2021; Habib et al., 2024; Huang et al., 2022), which can be readily scaled and adapted for integration into robotic systems designed for automated pest recognition tasks (Schmitz et al., 2022). Machine learning (ML) models are employed to identify and categorize pests, thereby reducing labor costs and facilitating the targeted application of chemicals (Domingues et al., 2022; Pattnaik & Parvathi, 2020). Smart or precision agriculture, characterized by the use of intelligent technologies and research on pest recognition, is experiencing rapid expansion (Kounani et al., 2024). Convolutional Neural Networks (CNNs) have demonstrated exceptional success in image classification tasks, particularly for identifying agricultural pests and diseases (Jia et al., 2020; Peng et al., 2023; Polin et al., 2024). Convolutional Neural Networks (CNNs) can learn discriminative features and patterns from extensive, annotated image datasets, enabling accurate and efficient pest classification. The classification of tomato pests using image analysis techniques has garnered considerable attention recently, driven by the growing demand for automated pest detection and management in the agricultural sector (Abekoon et al., 2024).

Convolutional neural network models have demonstrated effectiveness in agriculture, particularly for image processing and the detection of various agricultural targets, including fruits, leaves, weeds, pests, and crop rows. Researchers have investigated a diverse array of approaches and methodologies to tackle this challenge, leveraging advancements in computer vision (KY et al., 2024), machine learning (Mputu et al., 2024), and deep learning (Dharmasastha et al., 2022; Hong et al., 2020; Sardar et al., 2023). These models have consistently exhibited strong performance in addressing a myriad of crop-related challenges, including the classification of crop leaves (Hall et al., 2015), the classification of crop diseases (Mohanty et al., 2016), the assessment of crop yields (Rahnemoonfar & Sheppard, 2017), the recognition of crops (Lee et al., 2015), and the classification of crop types (Dyrmann et al., 2016). Rangarajan et al. proposed a pre-trained deep learning methodology for classifying tomato diseases (Rangarajan et al., 2018). The authors used images from the PlantVillage dataset, including diseased and healthy tomato leaves, as inputs to two deep learning-based networks, specifically AlexNet and VGG16. The model's effectiveness was evaluated by varying the number of images, minibatch sizes, and learning rates for weights and biases. The authors concluded that no discernible relationship existed between classification accuracy and the adjustments made to the AlexNet minibatch size. However, as the minibatch size increased, VGG16's accuracy declined. An advanced deep residual network approach was proposed by Yu et al. for diagnosing tomato pests, accurately identifying melon flies, cotton borers, green peach aphids, beet armyworms, taro insects, whiteflies, and spider mites (Yu et al., 2022). Wang et al. developed a sophisticated algorithm for real-time identification of pests and diseases affecting tomatoes in their natural environment (Wang et al., 2021). The proposed approach demonstrated improved detection performance when tested in actual tomato fields. Experimental findings indicated that this advanced method enhanced the capability to detect small and occluded objects. In 2020, Pattnaik and Parvathi proposed an advanced deep learning framework for recognizing tomato pests (Pattnaik & Parvathi, 2020). This innovative approach used pre-trained models, which were subsequently refined to effectively classify various tomato pests. By integrating Convolutional Neural Networks (CNNs) with transfer learning, the methodology achieved impressive accuracy in disease and pest identification. Xu et al. introduced a sophisticated, lightweight network for categorizing tomato pests and diseases, referred to as the Categorization Neural Network Architecture (CNNA) (Xu et al., 2023). The authors proposed implementing a multi-scale feature fusion module to enhance the model's ability to extract intricate features across diverse pest species and varying spot sizes. Furthermore, they implemented a global channel attention mechanism to enhance the model's sensitivity and precision in detecting pest characteristics. Experimental results showed that the fine-tuned lightweight models, such as MobileNetV3, MobileViT, and ShuffleNetV2, achieved lower classification accuracy than the proposed CNNA model. Fuents et al. proposed a refined filter bank framework for identifying pests and diseases in tomato plants, specifically addressing class imbalance and false positives (Fuentes, Yoon, et al., 2017). The proposed methodology effectively manages class imbalances and mitigates false positives from the bounding box generator by implementing a refinement filter bank framework, particularly in sparse-data settings. Sun et al. introduced an innovative approach that utilizes a lightweight neural network to identify tomato diseases and pests (Sun et al., 2024). Their research comprehensively examined the implications of dataset balance, data volume, and hyperparameter optimization to enhance model efficacy in diagnosing various tomato ailments and pest infestations. The results of their experiments substantiated the advantages of their proposed framework for tasks pertaining to the classification of tomato-related disorders.

Fuentes et al. proposed a sophisticated, real-time, deep learning-based detector for identifying tomato diseases and pests (Fuentes, Yoon, et al., 2017). The research conducted a comparative analysis of the performance of three prominent models: the Single-Shot Multibox Detector (SSD), the Region-based Fully Convolutional Network (R-FCN), and the Faster Region-based Convolutional Neural Network (Faster R-CNN). This evaluation used a benchmark dataset comprising tomato images, including nine distinct pest and disease types, while accounting for intra- and inter-class variations. The study meticulously examined the strengths and weaknesses of each model in the context of the pest classification task. Moreover, incorporating data augmentation techniques significantly enhanced classification performance. Liu and Wang introduced an advanced YOLOv3 architecture for identifying tomato diseases and pests, leveraging deep learning-based object detection (Liu & Wang, 2020). The researchers compiled a comprehensive dataset of pests and diseases affecting tomatoes in their natural environments. Implementing an image pyramid strategy significantly improved the YOLOv3 model's detection accuracy and processing speed by enabling multi-scale feature detection across the feature layers. The results of their study demonstrate that the refined YOLOv3 model outperforms traditional detection methods across performance metrics. Deng et al. devised an advanced methodology for detecting ten tea crop pests, employing a regional configuration technique to extract texture features akin to human vision, subsequently applying Support Vector Machine (SVM) classification (Deng et al., 2018). Although this innovative approach demonstrated precise pest detection in intricate environments, the inherently time-intensive preprocessing and manual feature extraction associated with machine learning techniques curtail their practical implementation. Leonardo et al. conducted a comprehensive evaluation of the efficacy of five deep learning architectures (VGG16, VGG19, InceptionV3, Xception, ResNet50) alongside nine machine learning algorithms (SVM, KNN, decision trees, MLP, and Naïve Bayes) for detecting pathogenic insects in fruit (Leonardo et al., 2018). By leveraging pretrained deep learning models on ImageNet, they found that integrating VGG16 with an SVM achieved the highest accuracy of 68.95%. Moreover, they advocated for implementing distinct, enhanced validation datasets to facilitate real-time pest detection. In a supplementary investigation, Thenmozhi and Reddy meticulously identified and classified various insect pests utilizing two datasets and databases from the NBAIR (xie1, xie2) (Thenmozhi & Reddy, 2019). They proposed an advanced model leveraging pre-existing neural networks, specifically employing VGG16, VGG19, GoogLeNet, ResNet, and AlexNet, and accomplished impressive accuracy rates of 96.75%, 97.47%, and 95.95% across the three respective databases. This exemplifies the prowess of deep learning models in autonomously discerning critical features, thereby augmenting the efficacy of pest identification and classification. Despite the abundance of research addressing diseases affecting tomato leaves and the evaluation of product quality (Cheng et al., 2019; Fuentes, Yoon, et al., 2017; Fuentes et al., 2018; Shijie et al., 2017; Sun et al., 2018) there is a definitive paucity of investigations specifically aimed at the detection of tomato crop pests through advanced image processing techniques (Fuentes, Im, et al., 2017; Nieuwenhuizen et al., 2018). This scarcity suggests that the geographical context of their cultivation likely shapes the pest varieties affecting tomatoes.

Previous research has extensively employed convolutional neural network (CNN)-based techniques for pest detection and classification. Additionally, various deep learning architectures have been utilized either as feature extractors or as direct classifiers. However, most studies have focused on classification accuracy, while the balance among accuracy, training duration, and feature extraction time has received comparatively less attention. This consideration is of significant importance for practical pest detection applications. Pest recognition enables the identification and classification of pests on crops, helping farmers take appropriate measures to avert financial losses and mitigate food insecurity. In the present study, we focus on classifying the most prevalent tomato pests using Convolutional Neural Network (CNN) models. The primary objective is to develop a robust, automated system that accurately identifies common tomato pests from visual cues in images. This research underscores the pressing need to implement effective Integrated Pest Management (IPM) strategies in tomato cultivation to reduce crop losses and improve agricultural productivity. Recognizing tomato pests presents numerous challenges, including the diverse appearances of pests across life stages, fluctuations in lighting and environmental conditions, and occlusions and clutter in images. Our research significantly advances automated pest recognition systems in the agricultural sector. This innovation facilitates early pest identification and the implementation of proactive pest management strategies, thereby ensuring sustained tomato crop production and productivity. Below is a summary of the significant contributions of our study:

Classification of tomato pests utilizing Convolutional Neural Network (CNN) models: We selected and implemented four prominent deep learning models to classify eight pest species plus one healthy class. Furthermore, we employed transfer learning and data augmentation techniques to improve evaluation metrics, including accuracy, precision, recall, and F-score.

Integrate Deep Learning and Machine Learning Models: This study systematically analyzes the impact of integrating deep learning and machine learning methodologies on the classification of tomato pests, offering actionable insights regarding their convergence properties and classification efficacy. Furthermore, the Grad-Cam algorithm was employed to visualize the behavior of convolutional neural network (CNN) models.

Comparing the proposed models not only in terms of classification accuracy but also in terms of training time and feature extractor in order to find the right balance between accuracy and time.

MATERIALS AND METHODS

The classification of tomato pests through image analysis requires a comprehensive integration of multiple components, including data acquisition, preprocessing, model architecture design, training, and evaluation. Figure 1 provides an overview of the classification methodology for tomato pests, using machine learning and deep learning techniques. By adhering to the comprehensive framework for tomato pest classification, researchers and practitioners can engineer an advanced, automated system for the photographic identification and categorization of tomato pests.

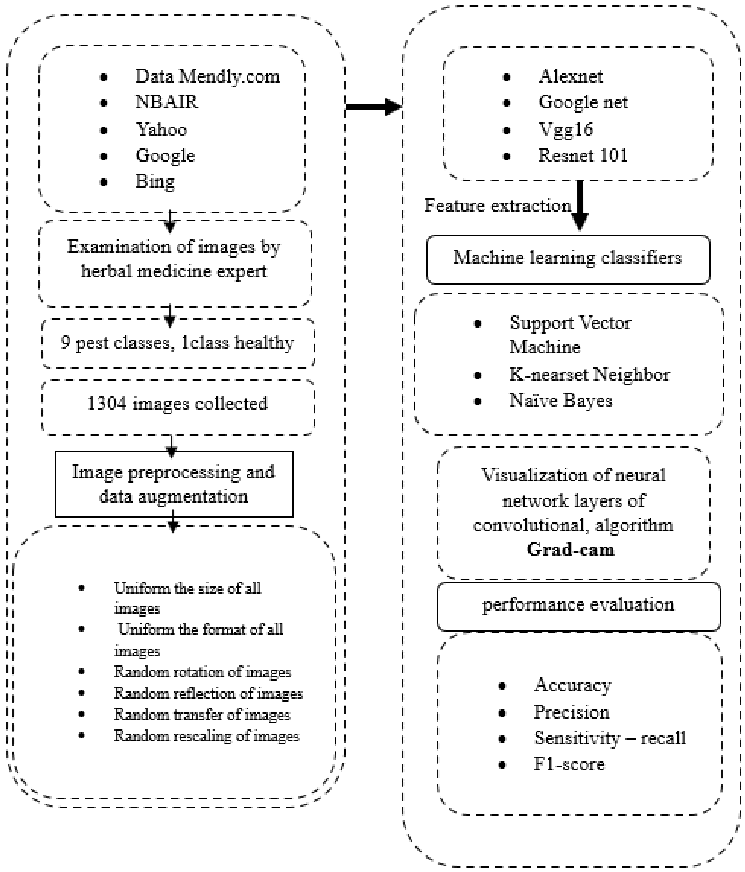


Figure 1. Visual Depiction of the Research Methodology

Image Database (Collection of Images)

The system initiates the process by aggregating a comprehensive and diverse dataset of images depicting tomato pests and healthy tomato samples. In total, 1,304 images were collected and used in this study, covering eight distinct tomato pest species and one healthy class, resulting in nine classes overall. Images for analysis were procured after the identification of the common and distinctive tomato pest species was confirmed by a plant protection specialist. Furthermore, the selected images were reviewed and validated at the image level by the same specialist to ensure that each image was correctly assigned to its corresponding class. The collection of these images involved using specialized image databases, such as Data Mendly.com and the National Bureau of Agricultural Insect Resources (NBAIR), as well as publicly available image results from prominent search engines, including Yahoo, Google, and Bing. As Data Mendly.com served as one of the main sources for the dataset, while supplementary images from NBAIR and search engines were used to increase dataset diversity and enrich classes with relatively few available images. The eight pests categories considered in this study are identified as follows:(1) Spider Mites (Tetranychus urticae)—TU, (2) Whitefly (Bemisia argentifolii)—BA, (3) Melon Fruit Fly (Zeugodacus cucurbitae)—ZA, (4) Melon Thrips (Thrips palmi)—TP, (5) Green Peach Aphid (Myzus persicae)—MP, (6) Tobacco Cutworm (Spodoptera litura)—SL, (7) Beet Armyworm (Spodoptera exigua)—SE, and (8) Fruit Borer (Helicoverpa armigera)—HA. Across diverse geographical regions, pests and insects exhibit considerable variation in species diversity, life cycles, and local nomenclature. Herbalists contend that pests which appear superficially identical may, in reality, represent distinct species. Using scientific nomenclature is essential for obtaining accurate and comprehensive search results when investigating pest species in digital databases and search engines. Owing to their remarkably similar morphology, the species Spodoptera exigua, Spodoptera litura, and Helicoverpa armigera, which are the central focus of this research, pose a risk of model confusion and erroneous classification. In addition to images of the eight pest species, this study included pest-free images to rigorously assess the efficacy of the classification models. Among the eight tomato pests examined, “Tetranychus urticae”, “Bemisia argentifolii”, and “Myzus persicae” are distinguished by their diminutive size and propensity to congregate in clusters on foliage. Detecting these minuscule pests in their natural habitat is arduous without magnification, complicating efforts to regulate their populations effectively. Figure 2 shows representative images of the pest species studied, while Table 1 offers a detailed overview of the eight pest classes, including one non-pest class. It includes their scientific names, image counts, and abbreviations.

Figure 2. Representative images of the pest species were gathered for analysis in this study

Table 1. Pest Class Abbreviations, Corresponding Scientific Names, and Number of Images for Each Class

Scientific name	Abbreviation	Number of images
Bemisia argentifolii	BA	60
Helicoverpa armigera	HA	108
Myzus persicae	MP	130
Spodoptera exigua	SE	75
Spodoptera litura	SL	96
Thrips palmi	TP	24
Tetranychus urticae	TU	74
Zeugodacus cucurbitae	ZC	42
Healthy	HL	695

Data Augmentation and Image Preprocessing

The image database contains diverse pest images captured from various angles, under different conditions, and in multiple locations, preparing them for Convolutional Neural Network (CNN) classification models. Two preprocessing stages were implemented. Initially, all images were standardized to the JPG format. Following this, the image dimensions were adjusted to align with the specifications of the CNN models and to enhance detection and classification precision. These steps were designed to improve dataset consistency, enhance model generalization, and optimize computational efficiency. Some variation in the number of images across classes was observed in the collected dataset. This imbalance may be related to natural differences in the occurrence, visibility, and image availability of pest species under different growth stages and field conditions. To mitigate the effects of class imbalance during training, data augmentation was applied as an important step in the preprocessing phase. Data augmentation was performed to enhance dataset diversity and improve the model's generalization capability. Specifically, the applied transformations included rotating images by +90 and -90 degrees, implementing horizontal and vertical flips, applying a cropping factor of 0.2, adjusting brightness to 1.2, and resizing images to 640 x 640 pixels. These augmented images replicate real-world variations, including differences in orientation, lighting conditions, and partial occlusions. This process increased the dataset size and helped reduce overfitting, thereby improving the model's performance on previously unseen data. The preprocessing steps established a reliable foundation for training the machine learning and deep learning models. By addressing issues such as varying image dimensions, limited dataset size, and categorical labels, the preprocessing pipeline ensured that the dataset was properly prepared for effective tomato pest classification. Following data augmentation, the training dataset was expanded to 6520 images. An illustration of this augmentation process is presented in Figure 3. To facilitate model training and evaluation, the collected images were divided into three subsets: training (70%), testing (20%), and validation (10%).

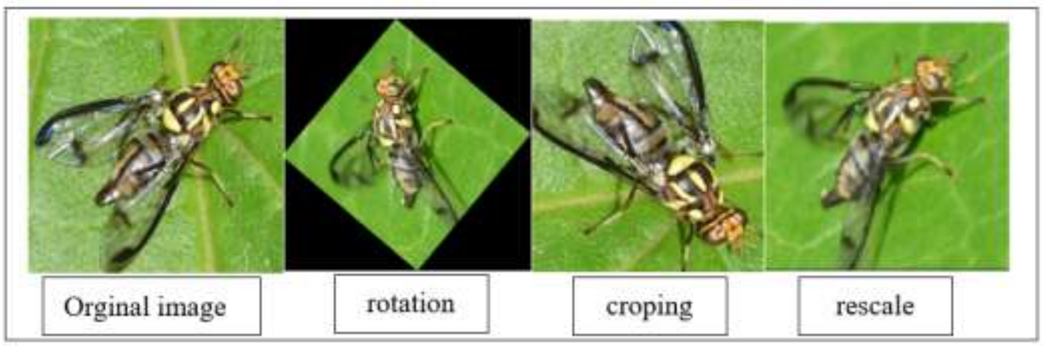


Figure 3. Illustration of the data augmentation technique

Convolutional Neural Network Models

In recent years, convolutional neural networks (CNNs) have gained significant recognition as a leading method for image classification. CNNs are widely utilized in diverse machine learning applications as a specialized branch of deep learning. The groundwork for neural networks was laid by the groundbreaking research of Hubel and Wiesel (Hubel & Wiesel, 1959). The classification of tomato pests using Convolutional Neural Networks (CNNs) marks a significant breakthrough in applying artificial intelligence to agricultural challenges. Convolutional Neural Networks (CNNs) have transformed pest identification and categorization by automatically learning and extracting complex image features. This innovation enables precise pest detection and improves management strategies. Training a convolutional neural network (CNN) on diverse datasets containing images of tomato pests enhances the model's ability to accurately distinguish among pest species and their developmental stages. This capability is essential for early pest detection and targeted intervention, which helps to reduce crop damage and enhance agricultural productivity. The Convolutional Neural Network (CNN) automates the labor-intensive and error-prone task of pest identification by analyzing distinct patterns, textures, and shapes found in pest images. Figure 4 presents the Convolutional Neural Network (CNN) model's architecture for classifying images of tomato pests. The Convolutional Neural Network (CNN) operates by computing features at each layer and passing them sequentially to subsequent layers through a process called forward propagation. Once the forward pass is complete, the model generates predictions, and the loss is calculated based on how much these predictions deviate from the actual labels. The backpropagation algorithm reduces loss by repeatedly updating the model's weights using gradient descent. This iterative optimization process enables Convolutional Neural Networks (CNNs) to progressively improve their predictions during training. The CNN architecture includes a single convolutional layer, which is crucial for extracting fundamental visual features such as edges and textures from input images. Following this layer is a batch normalization layer, which helps stabilize training and reduce the impact of covariate shift. Batch normalization improves a model's convergence speed and overall performance by normalizing the inputs to each layer. To further refine the features extracted, we incorporate a max-pooling layer. This layer reduces the spatial dimensions of the feature maps while preserving the most significant features. Additionally, it minimizes noise, ensuring that irrelevant image details do not negatively impact model performance. The processed feature maps are subsequently fed into dense layers that perform high-level reasoning and classification. The CNN models are equipped to effectively address the diverse and complex challenges in agricultural image data by combining convolutional layers, normalization, pooling, dense layers, and regularization techniques. This system lays the groundwork for advancing automated pest detection tools, contributing to sustainable and efficient pest management practices. In 2017, Ghazi et al. introduced AlexNet, a deep learning architecture that markedly outperformed traditional techniques (Ghazi et al., 2017). Since then, CNN models have undergone rapid advancements and are now employed across a range of applications. The remarkable efficacy of convolutional neural networks has enabled them to tackle a wide range of image classification challenges in the agricultural sector. Adhering to the conventional architecture of convolutional neural networks (CNN), illustrated in Figure 4, and the network comprises input, convolution, pooling, fully connected, and output layers. Activation functions are introduced to incorporate nonlinearity, enabling the network to discern complex patterns and approximate nonlinear functions. In essence, activation functions enable CNNs to learn intricate features from images, including edges, textures, and object components. The fully connected layer employs either a Softmax or ReLU activation function. Convolutional neural networks (CNNs) can autonomously extract features from images (Thenmozhi & Reddy, 2019). Specifically, the convolution layer serves as a filter, playing a pivotal role in feature extraction. By reducing the number and dimensionality of parameters in images and eliminating extraneous information, the pooling layer helps prevent overfitting in the network (Singh et al., 2020; Yu et al., 2014). To achieve precise pest classification into distinct categories, the Softmax activation function is utilized in the fully connected layer to extract and interpret intricate image information. The cross-entropy function is the cost function for the convolutional neural network models in this research. While traditional architectures employ ReLU in the hidden layers and Softmax in the output layer, this research diverges by using Softmax in the models' output layers.

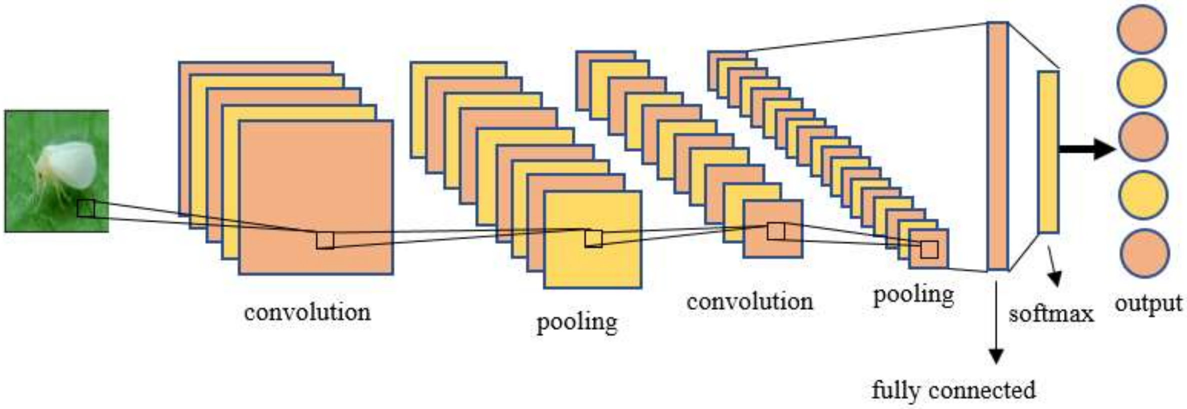


Figure 4. A Visual Representation of a Standard Convolutional Neural Network Architecture.

Transfer Learning Methodology

In their study, Ghazi et al. demonstrated the effectiveness of Convolutional Neural Networks (CNNs) in utilizing transfer learning through two distinct methods (Ghazi et al., 2017). The first method involves leveraging a pre-trained network to extract features that can be applied to a new problem, either by updating its weights or by using the already learned weights. This approach generates the required features prior to the final fully connected layer. In the second method, adjusting the output layer nodes to align with the number of classes relevant to the new issue is essential. This adjustment enables training and fine-tuning of the network using an updated database. The input data must also consist of images whose dimensions match those required by the pre-trained network. Both methods share this identical requirement. Convolutional neural network models require a large number of images. When sufficient data is available, deep learning approaches can enhance learning efficiency and reduce training time. Transfer learning allows the application of knowledge gained from one problem to a similar problem in a different context (Pan & Yang, 2009). This study evaluated several convolutional neural network models, including VGG16, ResNet101, GoogleNet, and AlexNet, using a custom image database we collected. These networks can be effectively trained on large-scale datasets such as ImageNet, enabling us to leverage pre-trained models to extract salient pest-related features. Subsequently, we adjusted the total number of output classes to align with the nine classes in our target dataset for this study. The model can be enhanced by using transfer learning, which is typically more efficient than training from scratch.

Experimental setup

In this study, eight categories of tomato crop pests were classified using pre-trained convolutional neural network models, employing transfer learning and fine-tuning. The pre-trained networks used in this study include VGG16, ResNet101, GoogleNet, and AlexNet, all of which have demonstrated strong performance on classification tasks. The hyperparameters of these networks have been adjusted based on methods outlined in prior literature (Coulibaly et al., 2019; Li et al., 2020; Thenmozhi & Reddy, 2019). In weight optimization, specifically in training neural networks, we employed Stochastic Gradient Descent with Momentum (SGDM), setting the momentum parameter to 0.9. For the classification layers of these networks, we applied the softmax activation function. The other learning parameters considered included an initial learning rate of 0.0001, a mini-batch size of 16, and a total of 150 epochs. The hyperparameter configurations, encompassing the learning rate, batch size, number of epochs, and optimizer settings, were determined based on widely adopted standards documented in prior research on transfer learning and image classification, as well as the requisites of the transfer learning framework and preliminary testing conducted during the experimental phase. To ensure equitable comparison, identical training parameters were applied consistently across all evaluated models.

Combination of deep learning and machine learning models (CNN+ML)

Convolutional Neural Networks (CNNs) provide the important advantage of automatically extracting salient information for pest classification. Unfortunately, their inability to accelerate training necessitates external computing units such as GPUs. The duration of the training process can range from one hour to several hours if the distinct compute units are not utilized. This is dependent on the depth of the network and the number of hidden layers. In previous studies, machine learning models were initially used to perform complex pre-processing tasks such as extracting color, shape, and texture data (Fuentes et al., 2018; Leonardo et al., 2018). Combining machine learning models with convolutional neural networks allows for feature extraction and classification. In this study, this approach avoids manual feature extraction, reducing training time and helping specialists and farmers more effectively. Research has shown that different studies employ varying feature extraction layers in convolutional neural networks. For instance, four layers of the VGG19 model were used in one study to extract features from images (Alhindi et al., 2018). In another study, features related to pests and leaf diseases of tomato crops were extracted from the penultimate fully connected layer of the VGG16 network (Shijie et al., 2017). In our study, the output of the last fully connected layer was classified using the SoftMax activation function. In this study, the desired features are extracted from the last fully connected layer of the mentioned networks, and for classification, these features are fed into machine learning models. The work process is shown in Figure 5. This study uses three commonly used machine learning classifiers: support vector machines (SVMs), k-nearest neighbors, and naïve Bayes.

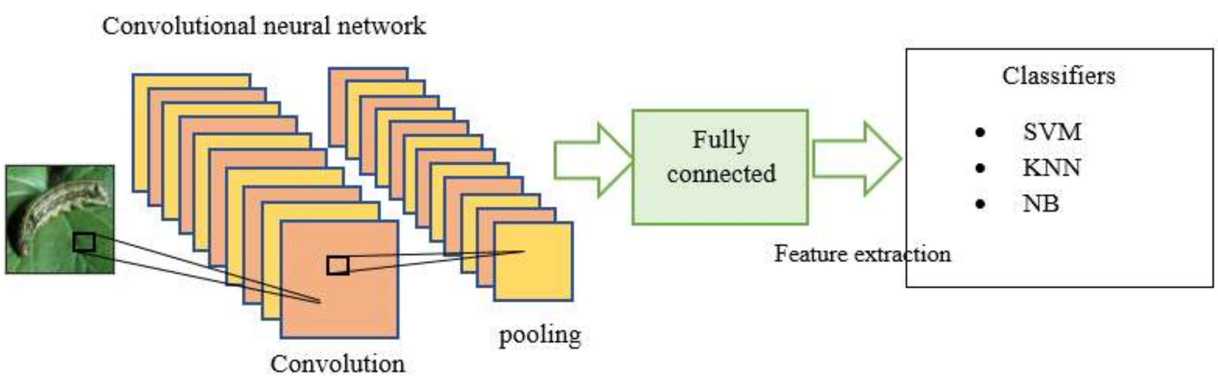


Figure 5. Combination of deep learning and machine learning models (CNN+ML).

Performance Assessment

The confusion matrix, accuracy, precision, sensitivity (recall), and F-score are five functional evaluation indices used to assess the performance of individual convolutional neural networks and combinations of convolutional neural networks with machine learning models.

Confusion Matrix

A confusion matrix can be used to calculate the number of false classifications and the accuracy of each classification made by a convolutional neural network when classifying a data collection. The following parameters can be obtained from a confusion matrix: True positive (TP), which refers to the positive tuples that are correctly tagged by the classifier. The number of these tuples is represented by TP; false positives (FP) are negative tuples that are falsely labeled as positive. The number of these tuples is represented by FP, True negative (TN). This term refers to negative tuples that the classifier correctly labels as negative. The number of these tuples is represented by TN, and false negatives (FN), Positive tuples that are wrongly labeled as negative. The number of these tuples is indicated by FN.

Accuracy

The accuracy of a classifier on the test data set is the percentage of images in the data set that are correctly labeled by the classifier, accuracy estimated with Eq. (1):

$$Accuracy = \frac{TP + TN}{P + N} \quad (1)$$

In Eq (1), class N is the total number of images in other classes, and P is the total number of images in the desired class; in other words, P+N is the total number of images in the test data set.

Precision

The precision of a classifier on the test data set is the percentage of images correctly labeled by the classifier, such that the label assigned by the classifier was indeed the true class label.

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

Sensitivity or Recall

The sensitivity of a classifier on test data is the percentage of correctly labeled images that are correctly classified, also known as recall.

$$Sensitivity\ or\ Recall = \frac{TP}{TP + FN} = \frac{TP}{P} \quad (3)$$

Accuracy and precision are inversely related; an increase in one may lead to a decrease in the other. Another solution for using the accuracy and precision measures is to combine them into a single, simpler measure called the F-SCORE evaluation index. This evaluation index is defined using Eq. (4).

$$F - SCORE = \frac{2 \times precision \times Recall}{precision + Recall} \quad (4)$$

RESULTS AND DISCUSSION

MATLAB R2022b has been used to perform processing and visualize the feature extraction process of convolutional neural networks in this study. The characteristics of the hardware used for image analysis and experimental tests are detailed in Table 2.

Table 2. Hardware characteristics.

Specifications
Laptop Asus N53 Core i7 (8GB\750HDD\Windows10\2GB graphics) Equipped with NVIDIA GeForce GT 550M CUDA

Visualization techniques can be used to understand how a convolutional neural network correctly classifies a test image. One of the easiest ways to understand how a network behaves when encountering various image classification problems is to visualize its different layers, such as convolutional and activation layers. Most convolutional neural networks learn to recognize features such as color and edges in the first convolutional layer. Gradually, as the network gets deeper, convolutional neural networks learn to learn more complex features. As demonstrated in (Figure 6-B), in the initial channels of the activation layer, simple features such as edges and lines are shown. The networks reveal more distinct aspects of the images as the layers and channels deepen. Figure 6-B shows the visualization of Google's network activation layers. White pixels indicate strong positive activations and black pixels indicate strong negative activations. As you can see, the network is initially learning low-level features such as lines. In the first channel of this layer, you can see that the pest's body is more prominent than the other parts because of its distinct lines and color compared to the rest of the image. More complex visualization methods can be used to further investigate a network's behavior. The Grad-CAM approach can be used to identify the regions of an image that may be crucial for the network to learn in a classification task, as well as to understand how the network behaves. The Grad-CAM algorithm uses the classification score of the gradients with respect to the convolutional features learned by the network; the regions where these gradients are large are precisely where the final score depends more on the data. Typically, feature maps are extracted by the Grad-CAM technique from the final layer of the linearly rectified unit (ReLU) activation function. In this study, these maps are also extracted from the Relu layer. The Grad-CAM feature maps show that the head and body parts of the pest can play a decisive role in the Googlenet convolutional neural network's prediction of the pest's class. This issue can be understood by looking at the more intense red areas around the mentioned areas.

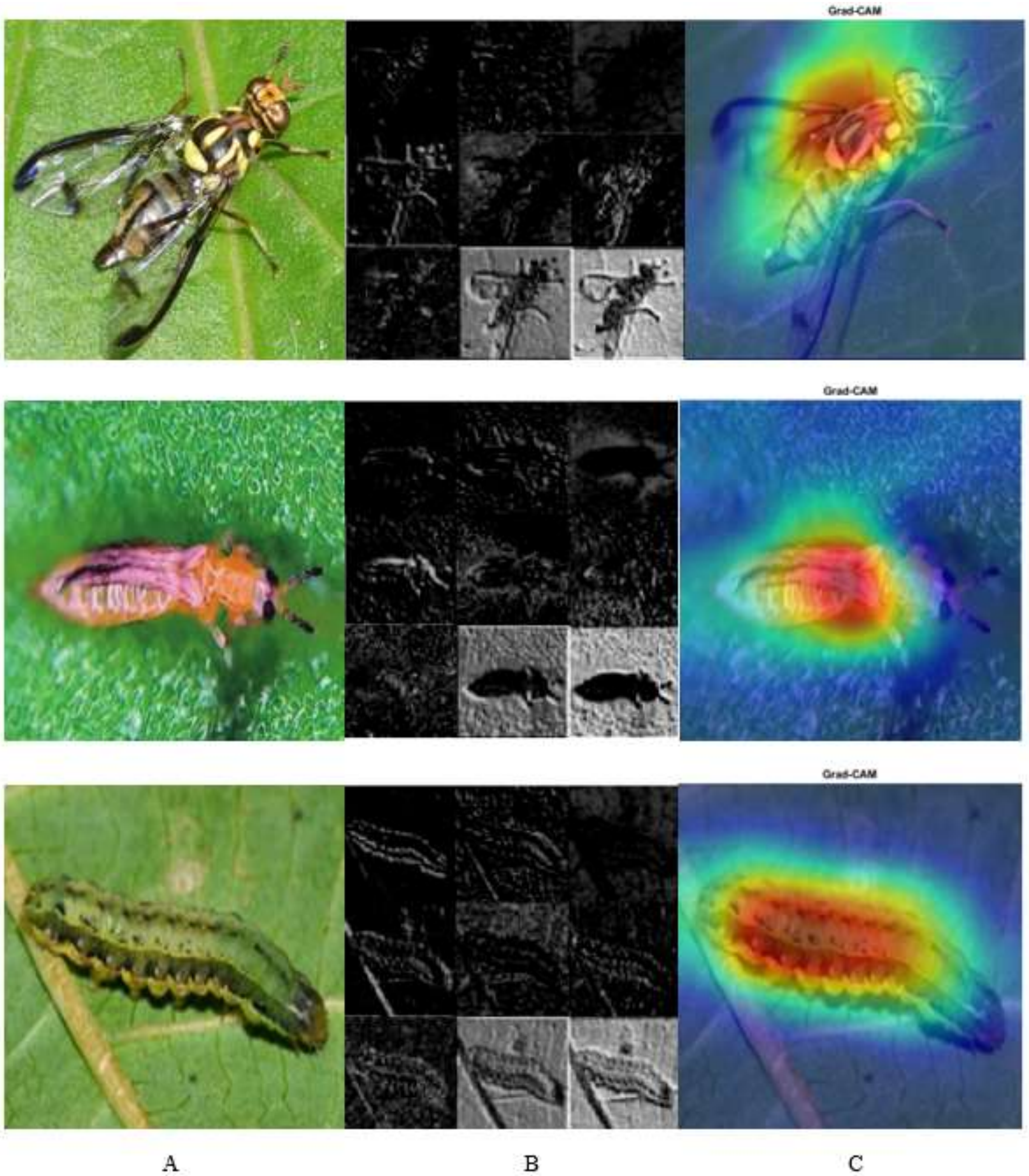


Figure 6. Visualization of feature extraction: A: Original image, B: Visualization of activation layer channels, C: Grad-CAM algorithm

Classification Results of Convolutional Neural Network Models

Tomato pest classification has been performed using transfer-learning-based convolutional neural network models. In the transfer learning approach, networks are pre-trained and re-applied to a new issue by fine-tuning some layers. Compared with training a network from scratch, this method will be easier and faster. The classification outputs and training times for each model are shown in Table 3.

Table 3. Training time and classification results for each of the models.

CNN models	Vgg16	Googlenet	Alexnet	Resnet101
Training time(ho:min:sec)*	06:05:23	01:08:06	32:20	04:52:44
Accuracy(%)	88.90	87.4	84.3	92.4
Precision (%)	85.47	82.55	76.8	83.67
Sensitivity or recall(%)	76.81	77.1	76	81.06
F-score(%)	79.5	78.56	75.5	82.02

The Resnet101 model achieves the highest accuracy, precision, sensitivity, and F-score among the models. (Table 3) displays the results obtained from the test data set. This good performance can be attributed to the network's deeper architecture, with 101 hidden layers, and its ability to extract complex features. For the Google Net and VGG16 models, the accuracies are 87.4% and 88.9%, respectively, a difference of 1.5%. According to the training times of 01':08":06'" and 06':05":23'" for the Google Net and VGG16 models, respectively, in scenarios where lower accuracy is acceptable but training time is shorter, Google Net is the better alternative between the two models. Among the four convolutional neural network models, the AlexNet model has the lowest recorded values for precision, accuracy, sensitivity, and F-score. The simplicity of the network is due to the use of a limited number of hidden layers, and also the inability to extract Suitable features for classification can be considered as one of the reasons for this poor performance.

Classification Results of Deep Learning + Machine Learning Approach

After features were extracted using convolutional neural network models, machine learning models were used to classify them. (Table 4) displays the training time for classification and the times required to extract features for each combination.

Table 4. Recorded times for training and feature extraction in every combination (minutes)

CNN+ML	Features extraction time	Training time
Googlenet+KNN	02:24	02:33
Alexnet+KNN	01:25	01:35
Resnet101+KNN	05:41	05:49
Vgg16+KNN	32:02	35:28
Googlenet+Svm	02:47	05:26
Alexnet+Svm	01:57	02:35
Resnet101+Svm	06:43	08:23
Vgg16+Svm	32:05	36:48
Googlenet+NB	03:47	06:47
Alexnet+NB	02:36	3:47
Resnet101+NB	07:48	10:56
Vgg16+NB	35:45	39:48

According to Table 4, the AlexNet+ML combinations have the shortest feature-extraction and training times compared with other models. This efficiency makes them particularly appealing for applications where speed is crucial. The network's simplicity and the small number of hidden layers could be the primary causes for the reduction in training time and feature extraction. After the Alexnet+ML combination, the Googlenet+ML, Resnet101+ML, and Vgg16+ML combinations have the lowest and highest feature-extraction and training-execution times, respectively. As we observe, as network depth increases, there is a corresponding rise in both feature extraction and training execution times. This trend highlights the trade-off between model complexity and computational efficiency. Classification results and evaluation indices for each combination of deep learning and machine learning are demonstrated in Table 5.

Table 5. Classification results and evaluation indices for each combination of deep learning and machine learning (percentage)

CNN models	Criterion	Googlenet	Alexnet	Vgg16	Resnet101
KNN	Accuracy	78.5	66.03	76.06	79.07
	Precision	74.13	72.03	71.41	77.93
	Sensitivity	60.62	49.85	57.63	63.81
	F-score	64.50	50.84	62.70	65.40
SVM	Accuracy	85.1	82.04	86.02	87.04
	Precision	82.39	75.28	81.02	83.67
	Sensitivity	70.24	70.03	72.64	79.63
	F-score	73.99	72.00	74.65	80.83
NB	Accuracy	79.06	65.04	78.03	81.35
	Precision	76.35	73.25	73.41	79.41
	Sensitivity	63.25	53.04	59.37	64.31
	F-score	66.85	55.17	66.38	69.17

The confusion matrix for the Resnet101+Svm combination, which achieves the highest performance across the evaluation indices among the combinations, is shown in Figure 7.

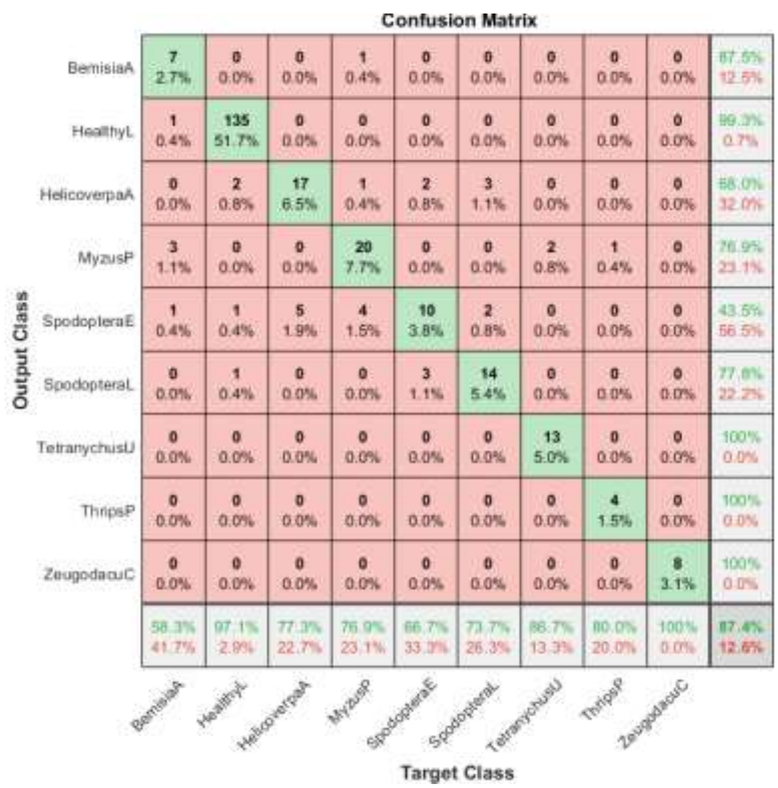


Figure 7. Confusion matrix obtained from Resnet101+Svm combination

The ResNet101+SVM model combination achieved the highest performance across the evaluation indicators, as shown in Table 5. This suggests that this combination effectively balances accuracy and efficiency in the evaluation metrics. For this combination, the accuracy, precision, sensitivity, and F-score were 87.4%, 83.67%, 79.63%, and 83.80%, respectively. Additionally, this combination demonstrated the shortest time for feature extraction and training, ranking only behind the AlexNet+ML and GoogleNet+ML combinations, as shown in Table 4. The utilization of 101 hidden layers in this network allows it to effectively extract complex features from images, contributing to its high performance. Also, the use of successive stacks of convolution and pooling layers and the Softmax activation function in the output layer may be among the reasons for the success of this combination. The performance of two other combinations of this model, i.e., Google Net+Svm and vgg16+Svm, is good, but the vgg16+Svm combination is not recommended due to its long feature extraction and training times. The results presented in (Table 4) and (Table 5) show that the classification performance for the Alexnet convolutional neural network model alone and also its combination with machine learning models does not seem very favorable. The results in Table 5 also show that, for all combinations of convolutional neural networks with a naïve Bayesian classifier, the accuracy, precision, sensitivity, and F-score indices have performed better than those of the CNN+KNN combinations. In general, from (Table 4) and (Table 5) regarding the training time, it can be concluded that the use of the CNN+ML combination can greatly reduce the training and feature extraction time. For example, the training time required for the VGG16 network alone is 06':05":23"', which is reduced to 39":48"' when combined with machine learning models such as naïve Bayesian.

Discussion

In previous related studies, various image databases have been employed for identification and classification. These databases provide diverse datasets that enhance the robustness of the models being evaluated. In particular, researchers have gathered images and developed image databases using various methods. For instance, some have assembled these databases by conducting Internet searches, while others have captured images directly in farm environments. These approaches help in creating diverse datasets tailored to specific research needs. Using the techniques outlined in the data collection section, we have assembled a database of images of tomato pests for this study, and the gathered image data was used to assess the chosen models. We've utilized several convolutional neural network models to achieve high classification accuracy. On the other hand, for devices with constrained computational capabilities, these convolutional neural networks require substantial processing power and lengthy training times. We have also combined CNN and ML models in this study. The combination of ResNet101+SVM demonstrated the best performance, achieving an accuracy of 87.4% according to the experiment's results. The bar chart in Figure 8 illustrates the F-score evaluation index for each convolutional neural network model. Figure 9 illustrates its combination with machine learning models for each pest class.

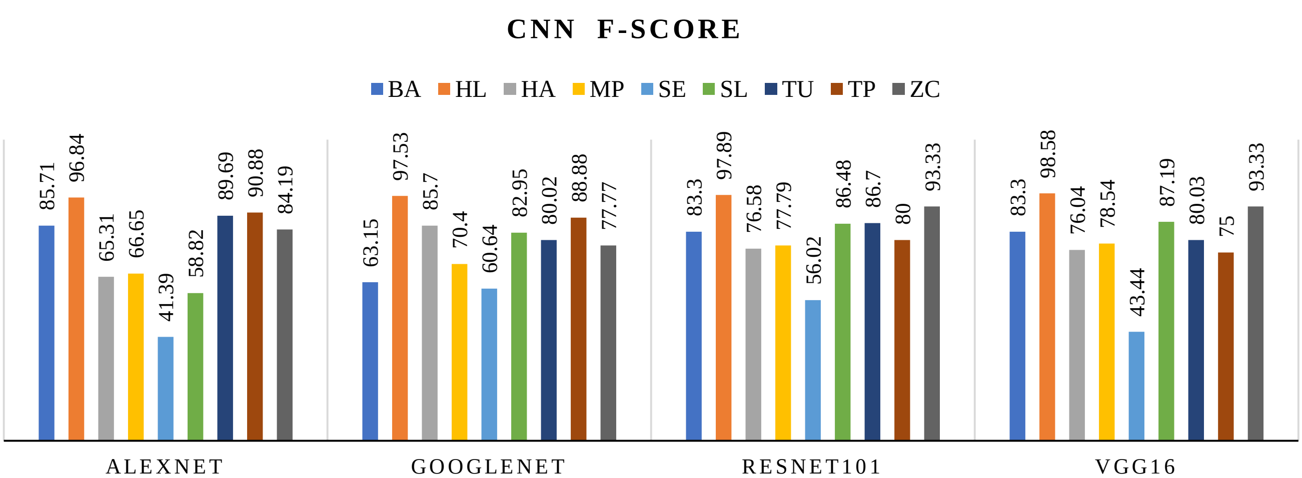
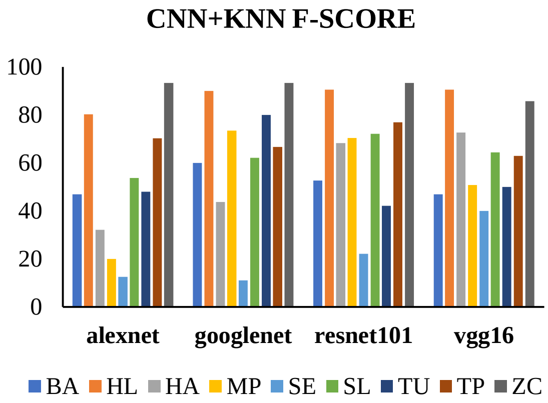
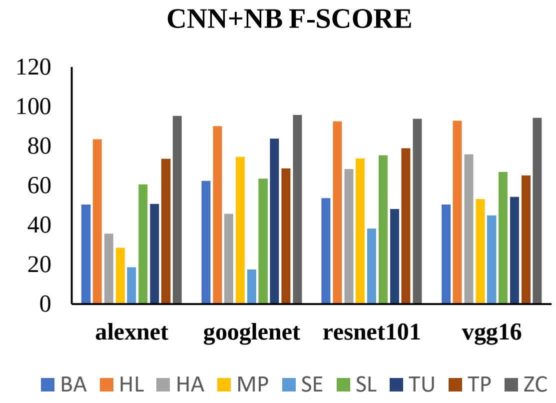


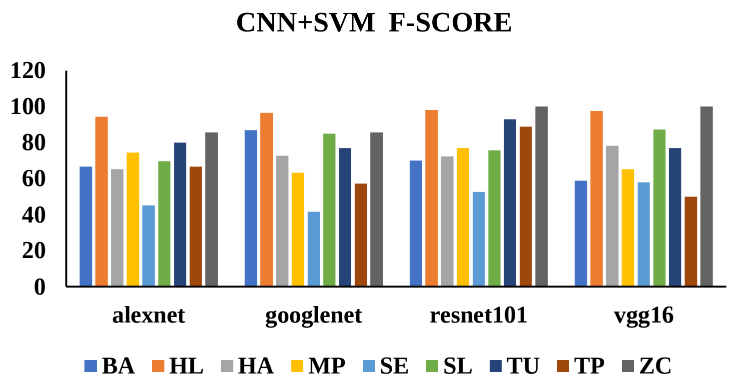
Figure 8. F-score evaluation index for each of the CNN models for each of the pest classes





A

B



C

Figure 9. A: CNN+Naïve Bayes B: CNN+K-nearest neighbor C: CNN+Support Vector Machine

The only pest class to achieve a perfect score on the F-score evaluation index is Zeugodacus cucurbitae. This outcome is illustrated in the diagram shown in (Figure 8-C), specifically for combining ResNet101 and VGG16 with SVM. From the analysis of the images related to this pest class, we have concluded that they are remarkably clear and contain very few irrelevant objects. This clarity likely contributes to the model's high performance in identifying Zeugodacus cucurbitae. The F-score evaluation metrics for this class in all three combinations of CNN+ML and convolutional neural networks seem very favorable. The performance of the classifiers is satisfactory for the Health Leaf pest-free class. All three combinations of CNN+ML and convolutional neural network models separately recorded values exceeding 90% for this class, except for two cases. The large number of educational images and the low complexity of the images in terms of irrelevant objects can be considered among the reasons for this performance. The F-score evaluation metrics in all three combinations of CNN+ML and convolutional neural network models, individually for the Spodoptera exigua pest class, recorded the lowest values, which does not seem favorable at all. The great similarity of this pest class with the Spodoptera litura and Helicoverpa armigera classes, as well as the excessive complexity of the images in this class, can be considered among the reasons for the unfavorable performance of the classifiers in dealing with this class. The Thrips palmi pest class, in which there are only 24 images, is well recognized and classified by the classifiers, so that the F-score evaluation metric for most of the models in this class is reported to be above 75%. In addition, the classifiers have shown an acceptable performance in recognizing and classifying these classes when faced with Myzus persicae, Tetranychus urticae, and Thrips palmi pest classes, which include pests with small bodies and are very similar to each other. So the F-score evaluation metric for these classes is reported above 75% in most classifiers. CNN models have been extensively employed in previous research to classify or find similar pests. In 2018, Nieuwenhuizen et al. classified tomato pests in the European region using CNN models; in their research, they achieved 87.40% accuracy (Nieuwenhuizen et al., 2018). However, tomato crops planted in different regions of the world are attacked by various pests. In this study, we have investigated eight common pests of tomato crops. Based on previous studies, we observe that the models proposed by Thenmozhi and Alves have the best classification performance with an accuracy of 95% (Alves et al., 2020; Thenmozhi & Reddy, 2019). In addition, the models used in this study have been used in other research as well, which has shown high accuracy in classification problems. For example, the VGG16 model in Cheng et al.'s research with 95.33% accuracy and the ResNet101 model in Cheng et al.'s research with 98.67% accuracy can be mentioned (Cheng, Zhang, Wu, et al., 2017; Cheng, Zhang, Chen, et al., 2017). When comparing the results of this study with previous research, it is essential to acknowledge that studies often use different datasets, pest species, environmental conditions, and evaluation protocols. Such variations can substantially influence the reported performance metrics of classification models. Accordingly, direct comparison of accuracy figures across different studies should be approached with caution. The comparisons provided herein aim to serve as a general reference for the performance of the proposed methodology, rather than as an exact or definitive benchmark.

CONCLUSIONS

The study we conducted included three additional evaluation metrics in addition to accuracy, whereas all the previously mentioned studies only employed one evaluation metric, which is accuracy. In this study, we evaluated the classification accuracy of various pest classes and analyzed the models' advantages and disadvantages from multiple perspectives. In this study, parameters are set based on previous literature, and the results demonstrated that the lower the initial learning rate is set, the lower the error rate. Furthermore, lengthier training time can result in more effective learning, but they also take a lot more time. In future research, in addition to the approaches used in this study, optimization processes can be utilized to improve the parameters of the convolutional neural network. In addition to using the SGDM training approach to update weights, other training approaches such as RMSprop, Adam, and Adadelta can also be used. Combining two convolutional neural networks, such as ResNet and GoogLeNet, can also be suggested to increase accuracy and improve classification.

DATA AVAILABILITY

The dataset generated during and/or analyzed during the current study is available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The code generated during and/or analyzed during the current study is available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contribution of the Biosystem Mechanical Engineering Department and the Computer Engineering Department of the Shahid Bahonar University of Kerman.

FUNDING

This research has not received any funded support.

CONFLICT OF INTEREST

The authors declare no conflict of interest in relation to this work or its publication.

REFERENCES

Abekoon, T., Sajindra, H., Jayakody, J., Samarakoon, E., & Rathnayake, U. (2024). Image processing techniques to identify tomato quality under market conditions. *Smart Agricultural Technology*, *7*, 100433. https://doi.org/10.1016/j.atech.2024.100433

Altuntaş, Y., & Kocamaz, F. (2021). Deep feature extraction for detection of tomato plant diseases and pests based on leaf images. *Celal Bayar University Journal of Science*, *17*(2), 145–157. https://doi.org/10.18466/cbayarfbe.812375

Alves, A. N., Souza, W. S., & Borges, D. L. (2020). Cotton pests classification in field-based images using deep residual networks. *Computers and Electronics in Agriculture*, *174*, 105488. https://doi.org/10.1016/j.compag.2020.105488

Amirahmadi, E., Ghorbani, M., Moudrý, J., Konvalina, P., & Kopecký, M. (2023). Impacts of environmental factors and nutrients management on tomato grown under controlled and open field conditions. *Agronomy*, *13*(3), 916. https://doi.org/10.3390/agronomy13030916

Bergougnoux, V. (2014). The history of tomato: from domestication to biopharming. *Biotechnology advances*, *32*(1), 170–189. https://doi.org/10.1016/j.biotechadv.2013.11.003

Cheng, H.-H., Ke, Y.-L., Lin, C.-P., Huang, J.-H., Chen, S.-F., & Kuo, Y.-F. (2019). Identifying and Localizing the Disease Spots of Late Blight on Tomato Leaves Using Deep Convolutional Neural Networks. 2019 ASABE Annual International Meeting (p. 1). American Society of Agricultural and Biological Engineers. https://doi.org/10.13031/aim.201900445

Cheng, X., Zhang, Y.-H., Wu, Y.-Z., & Yue, Y. (2017). Agricultural pests tracking and identification in video surveillance based on deep learning. International conference on intelligent computing (pp. 58-70). Cham: Springer International Publishing.

Cheng, X., Zhang, Y., Chen, Y., Wu, Y., & Yue, Y. (2017). Pest identification via deep residual learning in complex background. *Computers and Electronics in Agriculture*, *141*, 351–356. https://doi.org/10.1016/j.compag.2017.08.005

Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., & Traore, D. (2019). Deep neural networks with transfer learning in millet crop images. *Computers in industry*, *108*, 115–120. https://doi.org/10.1016/j.compind.2019.02.003

Deng, L., Wang, Y., Han, Z., & Yu, R. (2018). Research on insect pest image detection and recognition based on bio-inspired methods. *Biosystems Engineering*, *169*, 139–148. https://doi.org/10.1016/j.biosystemseng.2018.02.008

Depenbusch, L., Sequeros, T., Schreinemachers, P., Sharif, M., Mannamparambath, K., Uddin, N., & Hanson, P. (2025). Tomato pests and diseases in Bangladesh and India: farmers' management and potential economic gains from insect resistant varieties and integrated pest management. *International Journal of Pest Management*, *71*(4), 552–566. https://doi.org/10.1080/09670874.2023.2252760

Dharmasastha, K., Banu, K. S., Kalaichevlan, G., Lincy, B., & Tripathy, B. (2022). Classification of pest in tomato plants using CNN. International Conference on Metaheuristics in Software Engineering and its Application (pp. 56-64). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-11713-8_6

Domingues, T., Brandão, T., & Ferreira, J. C. (2022). Machine learning for detection and prediction of crop diseases and pests: A comprehensive survey. *Agriculture*, *12*(9), 1350. https://doi.org/10.3390/agriculture12091350

Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant species classification using deep convolutional neural network. *Biosystems Engineering*, *151*, 72–80. https://doi.org/10.1016/j.biosystemseng.2016.08.024

Fuentes, A., Im, D. H., Yoon, S., & Park, D. S. (2017, May). Spectral analysis of CNN for tomato disease identification. In International Conference on Artificial Intelligence and Soft Computing (pp. 40-51). Cham: Springer International Publishing**.** https://doi.org/10.1007/978-3-319-59063-9_4

Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, *17*(9), 2022. https://doi.org/10.3390/s17092022

Fuentes, A. F., Yoon, S., Lee, J., & Park, D. S. (2018). High-performance deep neural network-based tomato plant diseases and pests diagnosis system with refinement filter bank. *Frontiers in plant science*, *9*, 1162. https://doi.org/10.3389/fpls.2018.01162

Ghazi, M. M., Yanikoglu, B., & Aptoula, E. (2017). Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing*, *235*, 228–235. https://doi.org/10.1016/j.neucom.2017.01.018

Habib, S., Ahmad, M., Haq, Y. U., Sana, R., Muneer, A., Waseem, M., Pathan, M. S., & Dev, S. (2024). Advancing taxonomic classification through deep learning: a robust artificial intelligence framework for species identification using natural images. *IEEE Access*, *12*, 146718–146732. https://doi.org/10.1109/ACCESS.2024.3450016

Hall, D., McCool, C., Dayoub, F., Sunderhauf, N., & Upcroft, B. (2015). Evaluation of features for leaf classification in challenging conditions. *2015 IEEE Winter conference on applications of computer vision* (pp. 797-804). IEEE. https://doi.org/10.1109/WACV.2015.111

Hong, H., Lin, J., & Huang, F. (2020). Tomato disease detection and classification by deep learning. *2020 International conference on big data, artificial intelligence and internet of things engineering (ICBAIE)* (pp. 25-29). IEEE. https://doi.org/10.1109/ICBAIE49996.2020.00012

Huang, M.-L., Chuang, T.-C., & Liao, Y.-C. (2022). Application of transfer learning and image augmentation technology for tomato pest identification. *Sustainable Computing: Informatics and Systems*, *33*, 100646. https://doi.org/10.1016/j.suscom.2021.100646

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, *148*(3), 574. https://doi.org/10.1113/jphysiol.1959.sp006308

Jia, S., Gao, H., & Hang, X. (2020). Tomato pests and diseases classification model based on optimized convolutional neural network. *Journal of Physics: Conference Series* (Vol. 1437, No. 1, p. 012052). IOP Publishing. https://doi.org/10.1088/1742-6596/1437/1/012052

Jin, X., Zhu, X., Ji, J., Li, M., Xie, X., & Zhao, B. (2024). Online diagnosis platform for tomato seedling diseases in greenhouse production. *International Journal of Agricultural and Biological Engineering*, *17*(1), 80–89. https://doi.org/10.25165/j.ijabe.20241701.8433

Kounani, A., Tsimpiris, A., & Varsamis, D. (2024). Transitioning to a Sustainable Smart Agriculture Using Deep Machine Learning Techniques: The Case of a Greek Project. In *An Agenda for Sustainable Development Research* (pp. 39-54). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-65909-6_3

KY, P. A., Kulkarni, T. P., Rani, N. S., & MA, S. (2024, May). A novel approach for tomato quality evaluation: Computer vision in maturity assessment and disease detection. In *2024 5th International Conference for Emerging Technology (INCET)* (pp. 1-8). IEEE.. https://doi.org/10.1109/INCET61516.2024.10593108

Lee, S. H., Chan, C. S., Wilkin, P., & Remagnino, P. (2015, September). Deep-plant: Plant identification with convolutional neural networks. In *2015 IEEE international conference on image processing (ICIP)* (pp. 452-456). IEEE. https://doi.org/10.1109/ICIP.2015.7350839

Leonardo, M. M., Carvalho, T. J., Rezende, E., Zucchi, R., & Faria, F. A. (2018, October). Deep feature-based classifiers for fruit fly identification (Diptera: Tephritidae). In *2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI)* (pp. 41-47). IEEE. https://doi.org/10.1109/SIBGRAPI.2018.00012

Li, Y., Wang, H., Dang, L. M., Sadeghi-Niaraki, A., & Moon, H. (2020). Crop pest recognition in natural scenes using convolutional neural networks. *Computers and Electronics in Agriculture*, *169*, 105174. https://doi.org/10.1016/j.compag.2019.105174

Liu, J., & Wang, X. (2020). Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network. *Frontiers in plant science*, *11*, 898. https://doi.org/10.3389/fpls.2020.00898

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, *7*, 1419. https://doi.org/10.3389/fpls.2016.01419

Mputu, H. S., Abdel-Mawgood, A., Shimada, A., & Sayed, M. S. (2024). Tomato quality classification based on transfer learning feature extraction and machine learning algorithm classifiers. *IEEE Access*, *12*, 8283–8295. https://doi.org/10.1109/ACCESS.2024.3352745

Nieuwenhuizen, A., Hemming, J., & Suh, H. (2018). Detection and classification of insects on stick-traps in a tomato crop using Faster R-CNN. The Netherlands Conference on Computer Vision (pp. 1-4).

Oerke, E.-C. (2006). Crop losses to pests. *The Journal of agricultural science*, *144*(1), 31–43. https://doi.org/10.1017/S0021859605005708

Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, *22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Pattnaik, G., & Parvathi, K. (2020). Automatic detection and classification of tomato pests using support vector machine based on HOG and LBP feature extraction technique. In *Progress in Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2019, Volume 2* (pp. 49–55). Springer. https://doi.org/10.1007/978-981-15-6353-9_5

Peng, H., Xu, H., Gao, Z., Zhou, Z., Tian, X., Deng, Q., He, H., & Xian, C. (2023). Crop pest image classification based on improved densely connected convolutional network. *Frontiers in Plant Science*, *14*, 1133060. https://doi.org/10.3389/fpls.2023.1133060

Polin, J. A., Hasan, N., Habib, M. T., Rahman, A., Vasha, Z. N., & Sharma, B. (2024). Tomato pest recognition using convolutional neural network in Bangladesh. *Bulletin of Electrical Engineering and Informatics*, *13*(1), 619–627. https://doi.org/10.11591/eei.v13i1.6073

Praharsha, C. H., Poulose, A., & Badgujar, C. (2024). Comprehensive investigation of machine learning and deep learning networks for identifying multispecies tomato insect images. *Sensors*, *24*(23), 7858. https://doi.org/10.3390/s24237858

Rahnemoonfar, M., & Sheppard, C. (2017). Deep count: fruit counting based on deep simulated learning. *Sensors*, *17*(4), 905. https://doi.org/10.3390/s17040905

Rangarajan, A. K., Purushothaman, R., & Ramesh, A. (2018). Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia computer science*, *133*, 1040–1047. https://doi.org/10.1016/j.procs.2018.07.070

Sardar, P., Ema, R. R., Kabir, S. S., Adnan, M. N., & Galib, S. M. (2023). Severity Stage Identification and Pest Detection of Tomato Disease Using Deep Learning. *Int. J. Comput.*, *22*(2), 191–201. https://doi.org/10.47839/ijc.22.2.3088

Schmitz, A., Badgujar, C., Mansur, H., Flippo, D., McCornack, B., & Sharda, A. (2022). Design of a reconfigurable crop scouting vehicle for row crop navigation: a proof-of-concept study. *Sensors*, *22*(16), 6203. https://doi.org/10.3390/s22166203

Shijie, J., Peiyi, J., & Siping, H. (2017). Automatic detection of tomato diseases and pests based on leaf images. 2017 Chinese automation congress (CAC) (pp. 2537-2510). IEEE. https://doi.org/10.1109/CAC.2017.8243388

Singh, P., Raj, P., & Namboodiri, V. P. (2020). EDS pooling layer. *Image and Vision Computing*, *98*, 103923. https://doi.org/10.1016/j.imavis.2020.103923

Sun, J., He, X., Ge, X., Wu, X., Shen, J., & Song, Y. (2018). Detection of key organs in tomato based on deep migration learning in a complex background. *Agriculture*, *8*(12), 196. https://doi.org/10.3390/agriculture8120196

Sun, L., Liang, K., Wang, Y., Zeng, W., Niu, X., & Jin, L. (2024). Diagnosis of tomato pests and diseases based on lightweight CNN model: L. Sun et al. *Soft Computing*, *28*(4), 3393–3413. https://doi.org/10.1007/s00500-023-08631-w

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

Wakil, W., Brust, G. E., & Perring, T. M. (2018). Tomato and management of associated arthropod pests: past, present, and future. *Sustainable management of arthropod pests of tomato*, 3–12. https://doi.org/10.1016/B978-0-12-802441-6.00001-2

Wang, X., Liu, J., & Zhu, X. (2021). Early real-time detection algorithm of tomato diseases and pests in the natural environment. *Plant Methods*, *17*(1), 43. https://doi.org/10.1186/s13007-021-00745-2

Xu, Y., Gao, Z., Zhai, Y., Wang, Q., Gao, Z., Xu, Z., & Zhou, Y. (2023). A CNNA-based lightweight multi-scale tomato pest and disease classification method. *Sustainability*, *15*(11), 8813. https://doi.org/10.3390/su15118813

Yu, D., Wang, H., Chen, P., & Wei, Z. (2014, October). Mixed pooling for convolutional neural networks. In *International conference on rough sets and knowledge technology* (pp. 364-375). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-11740-9_34

Yu, H., Liu, J., Chen, C., Heidari, A. A., Zhang, Q., & Chen, H. (2022). Optimized deep residual network system for diagnosing tomato pests. *Computers and Electronics in Agriculture*, *195*, 106805. https://doi.org/10.1016/j.compag.2022.106805