Vol.68: e25240501, 2025
https://doi.org/10.1590/1678-4324-2025240501
ISSN 1678-4324 Online Edition

BABT
Brazilian Archives of Biology and Technology

Article - Engineering, Technology and Techniques

An Effective Feature Extraction Method for Tomato Leafminer - *Tuta Absoluta* (Meyrick) (Lepidoptera: Gelechiidae) Classification

Tahsin Uygun[1*]
https://orcid.org/0000-0002-9625-9513

Serhat Kiliçarslan[2]
https://orcid.org/0000-0001-9483-4425

Cemil Közkurt[3]
https://orcid.org/0000-0003-1407-9867

Mehmet Metin Ozguven[4]
https://orcid.org/0000-0002-6421-4804

[1]Burdur Mehmet Akif Ersoy University, Department of Electrical-Electronics Engineering, Burdur, Türkiye; [2]Bandirma Onyedi Eylul University, Department of Software Engineering, Balıkesir, Türkiye; [3]Bandirma Onyedi Eylul University, Department of Computer Engineering, Balıkesir, Türkiye; [4]Ankara University, Department of Agricultural Machinery and Technologies Engineering, Ankara, Türkiye.

Editor-in-Chief: Alexandre Rasi Aoki
Associate Editor: Fabio Alessandro Guerra

Received: 09-Jul-2024; Accepted: 20-Dec-2024

*Correspondence: tuygun@mehmetakif.edu.tr; Tel.: +90-248-2132802 (T.U.).

HIGHLIGHTS

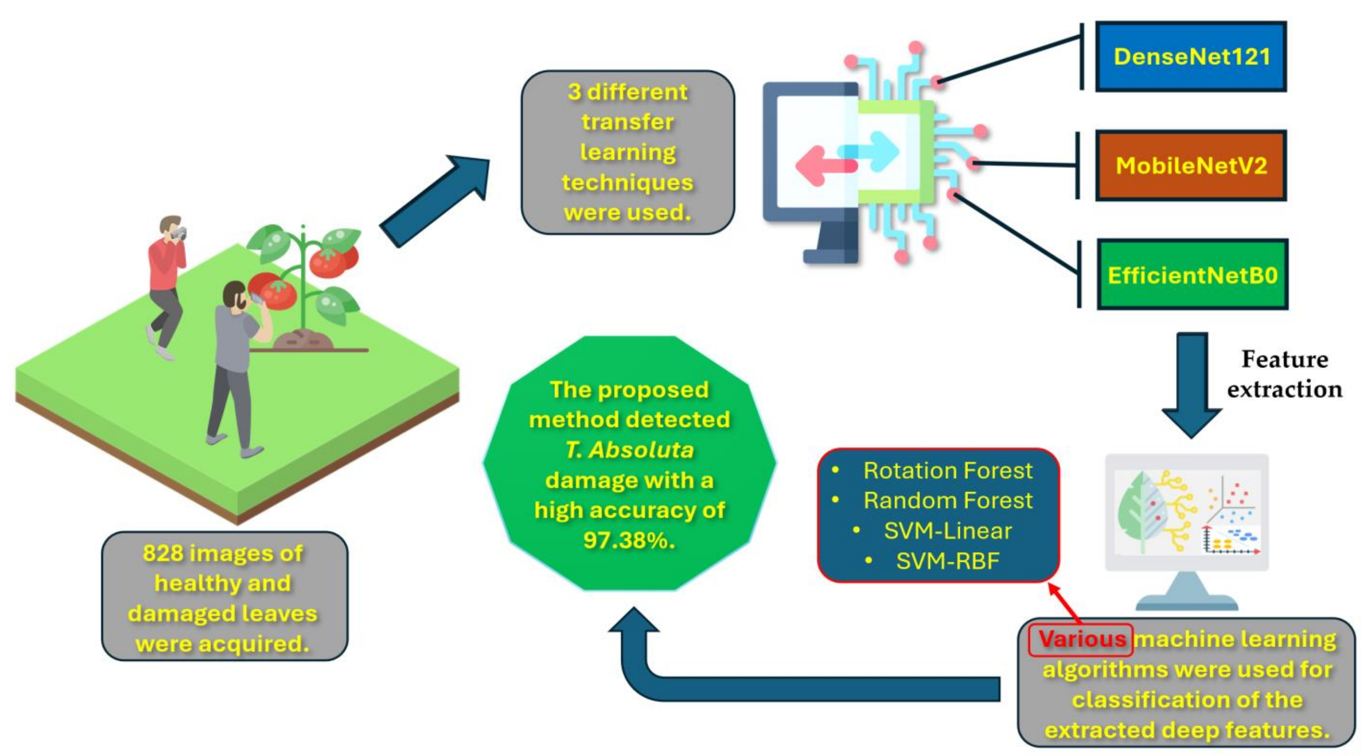
- A dataset consisting of healthy and damaged tomato leaves was generated.
- An effective method was improved to detect damage caused by *Tuta absoluta* pest on tomato leaves using a combination of transfer learning and feature extraction and machine learning approaches.
- In the study, feature extraction was performed with DenseNet121, MobileNetV2 and EfficientNetB0, and classification was performed with different machine learning algorithms.
- The method showed high performance with 97.83% accuracy rate in *Tuta absoluta* damage detection.

Abstract: Global warming caused by climate change causes some problems in agricultural production. One of these problems is the increase in various pest populations. This increase poses a serious threat to agricultural products and significantly negatively affects productivity and quality. Insecticides are commonly used to combat pests. However, most of the time, farmers' lack of knowledge in recognizing pests and understanding their effects results in incorrect and excessive spray applications. While excessive use of insecticides harms human health and environmental pollution, it also increases production costs, causes changes in the genetic structures of pests, causing them to become more resistant, and makes agricultural control difficult. Therefore, early detection of pests and their damage to the plant is extremely important. This study aims to develop an accurate and efficient method to detect damage caused by the tomato leaf miner, *Tuta absoluta*, on tomato leaves. A dataset comprising healthy and damaged tomato leaves was created.

Using a hybrid approach, features were extracted through Convolutional Neural Networks (CNNs) with transfer learning and classified using traditional machine learning techniques. Among the methods evaluated, SVM-Linear achieved the highest accuracy with 97.83%, outperforming other classifiers such as Random Forest with 96.14%, Rotation Forest with 95.89%, and SVM-RBF with 90.70%. These results highlight the potential of combining deep learning-based feature extraction with conventional machine learning for early pest detection. This approach offers a practical solution to reduce the misuse of insecticides and improve pest management strategies, contributing to sustainable agriculture.

Keywords: transfer learning; classification; feature extraction; feature selection; tomato leafminer.

GRAPHICAL ABSTRACT



INTRODUCTION

Many problems are encountered during agricultural activities due to the variety of products in agricultural production, the wide variety of cultivation areas and the labor-intensive cultivation. Technical possibilities offered by technology enable many agricultural processes to be facilitated, provide alternative solutions to some existing problems and especially by ensuring that production and applications are carried out correctly and on time, it makes significant contributions to the effective, efficient, and high quality of agricultural production. In addition, the use of technology has become necessary to meet the increasing food demand due to the increase in the world population and the decrease in agricultural areas [1]. Agricultural production can be adversely affected by many factors such as uncontrollable weather, climate, diseases, pests, drought, hail, flood, and economic crisis. In addition to expertise for efficient agricultural production, the most efficient production can be achieved if the growing conditions meet optimum demands. Ensuring optimum conditions can be difficult and costly. Therefore, for an efficient agricultural enterprise, it is necessary to determine the changes in each component accurately and in a timely manner and to decide on the applications considering possible interactions [2]. Plant disease and pest attacks cause significant economic and productivity losses in agricultural production worldwide [3-4]. As a result of climate change, the number of various plant diseases and pest species is increasing, and losses are gradually reaching serious levels. In order to minimize losses, plant diseases and pests must be combated [5]. *Tuta absoluta* (Meyrick) (Lepidoptera: Gelechiidae) pest, which originates from South America, has posed a major threat to tomato production globally and economically, with its rapid spread around the world [6-8]. When *T. absoluta* is not managed properly, it causes losses ranging from 80% to 100% in tomatoes [9]. Therefore, real-time and early identification of *T. absoluta* may play an important role in improving farmers' decisions in pest management [10]. The use of technological methods such as image processing and deep learning, which enable the rapid detection of disease and pest severity and development in the diagnosis of diseases and pests, offers advantages such as preventing loss of labor and time [11-14]. A key advantage of the deep learning method is that it can extract features on its own. In deep learning, the algorithm is not given any explicit instructions. Instead, it automatically searches through the data for features that correlate and then combines them to facilitate faster learning [15-17]. Because of its ability to handle big data, deep learning scales extremely well [18]. Feature extraction is a dimensionality reduction technique in which the original dataset, consisting of a set of raw variables, is reduced into sets of adaptable features for a new task, without changing them. Therefore, features must be extracted from the dataset to provide data to the classifier. For example, multiple feature extraction is used to classify tomato leaf diseases. Color histograms are extracted for color features, Hu-moments for shape features, and haralick features for texture features [19].

In this study, unlike previous studies, a new dataset was originally created for the detection of *T. absoluta* damage, deep features were extracted from these images using three different transfer learning techniques, DenseNet121, MobileNetV2 and EfficientNetB0, and then, to classify these deep features extracted, Rotation Forest, Random Forest, Support Vector Machine-Linear (SVM-Linear) and Support Vector Machine-RBF (SVM-RBF) machine learning algorithms were used. The objective of this study is to develop an accurate and efficient system for detecting *T. absoluta* damage on tomato leaves by combining advanced feature extraction methods with traditional machine learning algorithms. The proposed method is successful in capturing complex spatial relationships and patterns in images of tomato leaves and enabling the discernment of subtle pest damage indicators. The novelty of this study lies in its hybrid approach, combining deep learning-based feature extraction with traditional machine learning algorithms, as well as the creation of a real-world dataset specific to *T. absoluta* damage, which is rarely addressed in the literature. This innovation is of great importance as it increases the accuracy and efficiency of pest damage and disease detection compared to traditional methods. At the end of the study, damaged and healthy leaves were distinguished with a 97.83% accuracy rate in the classification processes carried out with machine learning approaches.

Related works

In agriculture, state-of-the-art (SOTA) artificial intelligence methods are used in many areas such as monitoring plant and fruit periods, yield estimation, species and quality classification, pest and disease detection, and proportional fertilization to increase the yield of agricultural products and reduce losses. In order to combat *T. absoluta*, the pest needs to be detected and classification, segmentation and object identification studies have been and are being carried out on image data as SOTA artificial intelligence technology.

The authors in [20] talk about how plant pests affect agriculture's bottom line, concentrating on *T. absoluta*, a significant pest of tomato plants. YOLOv5 was employed in the study to train on 1,200 altered pictures of tomato leaves infested with *T. absoluta*. The results showed a promising 80% classification accuracy for leaves and a 70–90% detection accuracy for *T. absoluta* larvae and galleries. According to the research, deep learning-based algorithms have great promise for the early identification of agricultural pests, which might lead to a decrease in the requirement for overuse of pesticides. *T. absoluta* and Leveillula taurica infected tomato plants were the subject of 263 multispectral (RGB and NIR) photos that the scientists [21] assembled into a single, unique dataset. To find and categorize the lesions on the leaves, they employed a Faster-RCNN object detector. In comparison to utilizing only RGB channels (mAP of 18.5%), their research showed that using both RGB and NIR channels enhanced the detection accuracy (mAP of 20.2%). This implies that NIR spectral bands may be useful in the identification of certain tomato plant diseases. Researchers [22] have created a novel AI-powered instrument that can identify *T. absoluta* infestations in tomato crops with 91.9% accuracy by using CNNs. With the use of the tool, farmers may be able to prevent crop losses by quickly detecting and treating pests. Two deep learning models, U-Net and Mask R-CNN are presented in a research [23] to help segment and identify the tomato leaf sections damaged by *T. absoluta*. These models accurately analyze tomato leaf photos recorded in the field at the pixel level using CNNs. With a mean Average Precision of 85.67%, the findings show that the Mask R-CNN model performs better in terms of precision than the U-Net model. With a Dice coefficient of 82.86% and an Intersection over Union (IoU) of 78.60%, the U-Net model still exhibits promising performance. Both models are useful resources for farmers and agricultural experts as they precisely identify and divide the regions where *T. absoluta* infestations are present. CNNs are being used in a study [10] to assess the extent of *T. absoluta* infestation in tomato plants. Using a dataset comprising both infected and healthy tomato leaves, the scientists trained four pre-trained CNN architectures. With an average accuracy of 87.2% in determining the extent of *T. absoluta* infestation, Inception-V3 produced the best findings. Additionally, the pre-trained models had no trouble differentiating between high, low, and no infestation of *T. absoluta*. A study [24] looked at the detection of T. absoluta infestations on tomato plants using object detection models in conjunction with ensemble approaches. In order to handle the intricacies of real-life plant health scenarios, the researchers employed real-world datasets collected in greenhouse and open-field conditions. Deep learning methods that showed potential in identifying *T. absoluta* damage were Faster R-CNN and RetinaNet. Early model assessments showed that detection performance was limited. With a 20% improvement in mean Average Precision (mAP), ensemble approaches such as Weighted Boxes Fusion (WBF) considerably increased detection accuracy. An effective technique for identifying peacock eye disease in olive leaves was discovered in a study [25]. An original dataset of healthy and disease leaves was constructed in the first step. The categorization of damaged and healthy leaves was then carried out using the transfer learning technique, which involved extracting deep features from this dataset using CNN models. Approximately 98.63% of extremely excellent findings were produced as a consequence of the trials.

MATERIAL AND METHODS

In this part of the article, *T. absoluta* pest and its damage to the tomato leaf, creation of the leaf dataset, use of machine learning methods and the obtained performance criteria are presented.

Tomato Leafminer - *Tuta absoluta* (Meyrick) (Lepidoptera: Gelechiidae)

Tomato cultivation faces various pests that increase the cost of production and lead to health and environmental risks due to pesticide use [21]. The most important of these pests is Tomato Leafminer (*Tuta absoluta*). *T. absoluta* causes damage to leaves, flowers, stems and fruits of tomato. [26]. This pest is a member of the order Lepidoptera, belonging to the family Gelechiidae [27]. In the egg stage (Figure 1a), it is 0.4 mm long and 0.2 mm wide, has a cylindrical structure and is cream to light yellow in colour [28]. The hatched young larvae (Figure 1b) enter the leaves, stems or tomato fruits where they feed and grow, forming distinct mines and galleries [29] This affects the plant's ability to photosynthesize, which reduces tomato yield. The pupa of the pest (Figure 1c) is 6 mm in size and has a color close to green in the first stage; then its color changes to light brown. The adult of the pest (Figure 1d) is 6 mm long, slender and has a wingspan of approximately 10 mm. *T. absoluta* can cause varying levels of damage to tomato leaves. If a high damage rate occurs (Figure 1g), the economic threshold value is exceeded and therefore it is critical to detect damage at an early stage. In order not to exceed this damage threshold, it is necessary to detect it at low severity and earlier damage stages, as given Figure 1f.

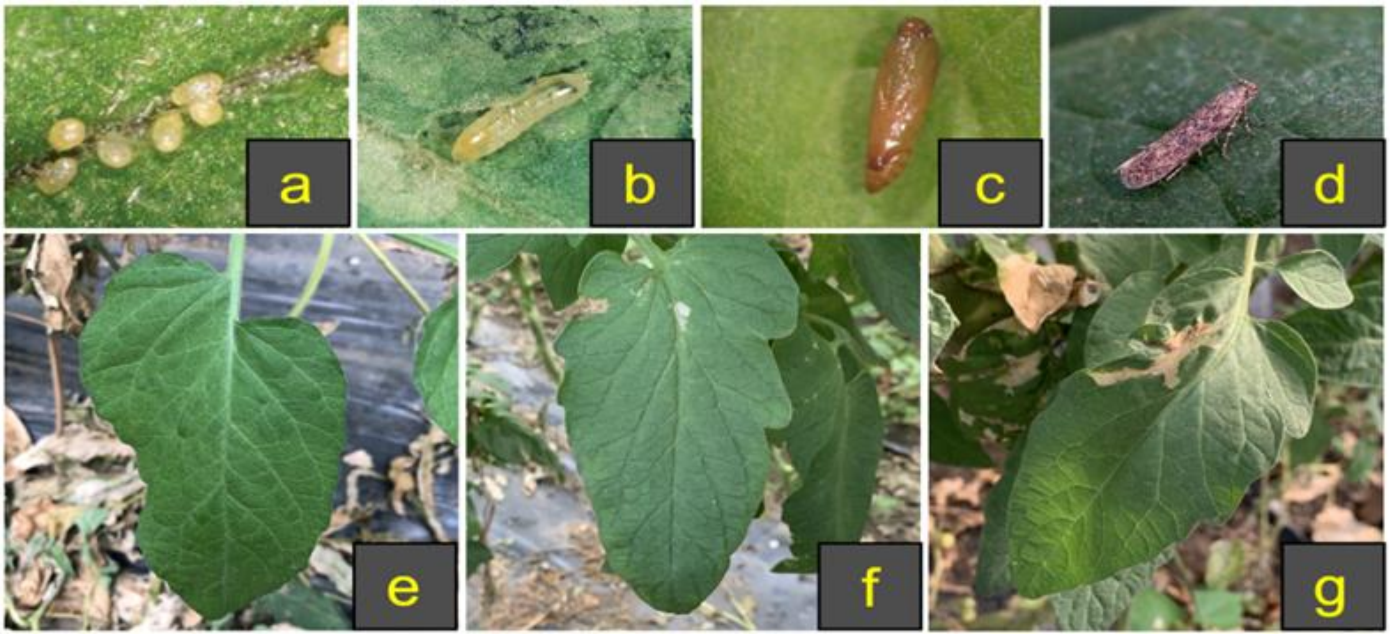


Figure 1. *T. absoluta*; (a) egg, (b) larva, (c) pupa, (d) adult and tomato leaf; (e) healty, (f) low damage, (g) high damage.

Dataset

To create the dataset used in the study, 828 tomato leaf images were obtained under greenhouse conditions on different dates (August 10 and September 7, 2023) from the village of Gümenek in the Central district of Tokat province. In this dataset, images are divided into two classes, as shown in the structure in Figure 2: damaged by *T. absoluta* and healthy tomato leaves.

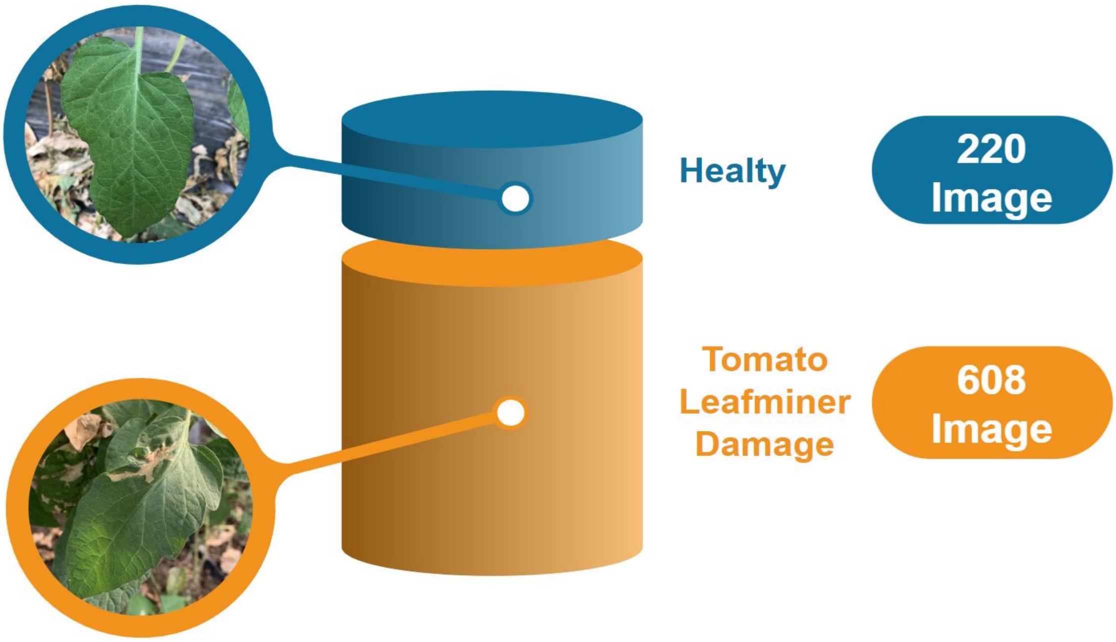


Figure 2. Dataset structure.

Image acquisition and preprocessing

Within the scope of this research, the dataset obtained from images with natural lighting conditions and complex background, as indicated in the structure in Figure 2, consists of 220 healthy and 608 damaged tomato leaves. When the flow chart in Figure 3 is examined, 2 smartphones (iPhone XS and iPhone 8, Apple Inc., California, USA) were used for data acquisition and image acquisition under greenhouse conditions. The raw dataset was created using 2x optical zoom with a resolution of 3024x3024 (1:1). Images were automatically transferred to the computer environment. Tomato leaf images were processed using a computer with a 12th Generation Intel Core i9-12900 (24 CPUs) processor running at 2.4 GHz and NVIDIA (Santa Clara, California, USA) RTX A4000 graphics.

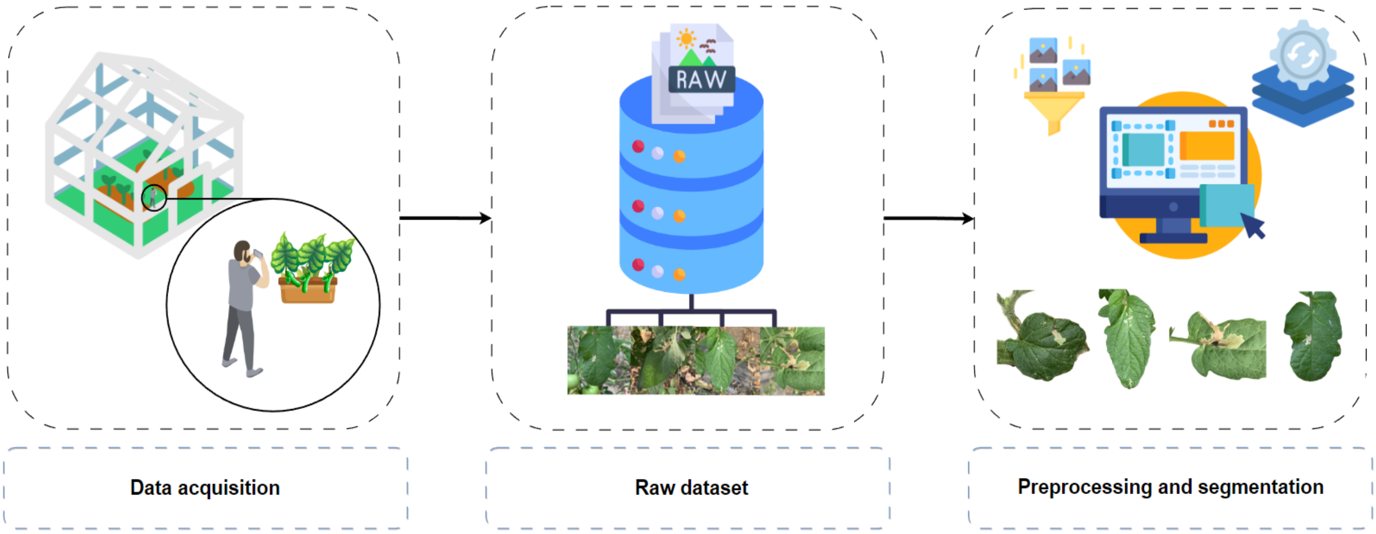


Figure 3. The data acquisition flowchart.

The image segmentation method is a three-stage process. Initially, the image data is read, and the acquired images undergo conversion from the RGB color space to the RGBA color space. This transformation includes adding a transparent background to the image, effectively eliminating unnecessary background details. In the second stage, a color-based segmentation mask is generated. In the collected dataset, accounting for the potential presence of white and occasionally blue lines in the background (addressed by employing the summation method for separation), binary masks are established in blue and white colors. These masks are subsequently combined using the bitwise OR operation. In the third stage, morphological operations are applied to the generated mask to close small gaps (noises). The result is a refined mask where regions consistent with the RGBA image mask are preserved, successfully isolating and filtering only the leaves [25].

Feature extraction

DenseNet121 [32], MobileNetV2, and EfficientNetB0, along with various other CNN architectures, have found widespread applications in image classification, demonstrating proficiency in tasks such as image recognition and object detection. The activation function is utilized in the convolutional layers to convert the learned kernels from preceding layers into feature maps, thereby contributing to the generation of output feature maps [30]. Convolutions involving multiple input feature maps can be amalgamated within each output feature map. These architectures, utilized in the current study, are acknowledged as the top three performers in terms of classification accuracy within the domain of deep learning. Each image encapsulates features determined by the fully connected layers of DenseNet121, MobileNetV2, and EfficientNetB0 before entering the classification stage. Subsequent to these layers, deep learning architectures progress to the classification process by incorporating softmax and classification layers. The features amassed are converted into deep features through machine learning techniques applied in the "fully connected" layer, thereby fulfilling the purpose of classification.

Transfer learning

Transfer learning is an approach that applies knowledge and qualities obtained while solving one problem to new, closely similar problems [31]. For example, the information and skills gained from classifying animal photographs can be applied to classifying floral images. Transfer learning is the process of improving one's learning of a new task by applying knowledge gained from a previous one. This strategy is more practical and effective than starting the learning process from scratch, especially when the availability of limited datasets makes training a model from scratch problematic [32].

DenseNet has transformed the deep learning model training landscape by efficiently addressing the vanishing gradient paradox, improving attribute reuse, and maximizing parameter consumption. Its substantial impact is visible in its outstanding performance across a wide range of computer vision challenges. The dense blocks (DenseBlock) are painstakingly formed using convolutional layers and are at the heart of DenseNet121. This architecture is made up of 121 interconnected convolutional layers that culminate in a final output layer with 1000 fully-connected units. This approach not only addresses the vanishing gradient difficulty, but it also maximizes attribute sharing, which contributes to its extraordinary efficacy in a variety of computer vision applications [33].

MobileNetV1 [34] a customized and efficient model designed for mobile and embedded devices, in 2017. This architecture emphasizes two stages of standard convolution: depth-decomposable convolution and point-decomposable convolution. The calculation of the bit product stays tightly related to the channels inside the MobileNetV1 architecture, ensuring a stable channel count across the model. The versatility of MobileNetV1 is particularly noteworthy, particularly in the configurable number of channels within the 11-convolution kernel. Because of its intrinsic flexibility, MobileNetV1 is able to strike a careful balance between computing efficiency and model correctness. MobileNetV2 [35], the subsequent version, offered novel features such as inverted residual structures and linear bottlenecks. Inverted residual structures give a solution to the limitation of in-depth convolution in modifying channel numbers. These configurations include an ascending convolution filter with 11 dimensions, a depth convolution filter with 33 dimensions, and a descending convolution filter with 11 dimensions. The activation function is meticulously interwoven into the inverted residual structural functions to accelerate the learning trajectory, mitigate gradient spread, and strengthen the model's global stability.

EfficientNet, a convolutional neural network architecture, was proposed in 2019 with a methodical approach to model scaling. This approach employs a composite coefficient within EfficientNet, allowing for the balanced tuning of network depth, width, and resolution, resulting in improved performance. All dimensions, including depth, width, and resolution, can be equally scaled using compound scaling. In this strategy, changing the size of the input image correlates to an expected need for more network layers to augment the receptive field and more channels to detect smaller patterns. The EfficientNetB0 variation has 5.3 million parameters and is specifically developed for use in mobile and embedded devices. Notably, the EfficientNetB0 core architecture is inspired by the bottleneck residual blocks inherent in the MobileNetV2 network [36]. The parameters of the transfer learning models used in the study are given in Table 1.

Table 1. Parameters of transfer learning algorithms used in the study.

Models	Parameters
DenseNet121	Activation function : ReLU Weights : imagenet input_shape : 224×224×3
MobileNetV2	Activation function : ReLU Weights : imagenet input_shape : 224×224×3
EfficientNetB0	Activation function : ReLU Weights : imagenet input_shape : 224×224×3

Machine learning methods

SVM-Linear

SVM is a discriminant hyperplane used to identify classes [37]. These selected hyperplanes help to maximize the distance to the nearest training points, called "margins". This feature is generally known to improve generalization capabilities. SVM uses a regularization parameter, C, which ensures the fit of outliers and allows for errors in the training set. It also allows classification using linear decision boundaries; this is called linear SVM. This classifier has been applied quite successfully to many brain-computer interface problems. Furthermore, kernel trickery is used to generate non-linear decision boundaries with less complexity. In EEG and image analysis, Gaussian or Radial Basis Function (RBF) kernels are usually the preferred kernel types. This type of SVM is also referred to as Gaussian SVM [38-39].

SVM-RBF

RBF can be thought of as a curve fitting approach in multidimensional space. RBF classifiers work by summing a set of radial basis functions weighted to approximate the desired function f. Therefore, the training of RBF involves the problem of finding a surface that best fits the training data in multidimensional space. RBF has a three-layer structure: input, hidden and output. The input layer is the layer connected to the input vector space. The output layer is the layer associated with the pattern classes. The training of the classifier takes place by determining the weights in the hidden layer that connects the input and output layers. The determination of the model parameters is related to the determination of each centre C_j and bandwidth σ_j for the activation functions used as Gaussian functions by the neurons in the hidden layer. RBF classifiers have a suitable format for multiple classification problems [40-41].

Random Forest

Random Forest (RF) can be defined as a collection of classifiers consisting of different types of trees. This approach is an ensemble learning algorithm for classification, regression and similar tasks that works by building a set of decision trees in the training phase [42]. Instead of branching each node by selecting the best split point among all variables, RF builds a tree with nodes that branch using the best among randomly selected variables. Each dataset is created by sampling from the original dataset and then trees are constructed using random feature selection [43]. In regression tasks, the predicted values of each tree are calculated and averaged, resulting in a prediction. Like k-nearest neighbor (k-NN) and pure Bayes-based algorithms, RF is a machine learning algorithm that has become popular due to its simplicity and high classification performance.

Rotation Forest

Rotation Forest (RTF) is a powerful ensemble classifier with impressive performance. This algorithm constructs the transformation matrix using Principal Component Analysis (PCA) and utilizes this matrix to extract new features [44]. The RTF classification algorithm is a linear transformation technique that offers a new performance angle by integrating a different domain [45]. The RTF algorithm follows the basic principles of the RF algorithm in classification, creating multiple trees. However, it differs from RF in that it uses an alternative feature space such as PCA to construct the data set. On training data sets defined by this feature space, RTF generates a large number of decision trees (DTs). To train these trees during training, the RTF algorithm subdivides the training dataset into subgroups and performs feature extraction using the feature space selected for each subgroup. [46] stated that thanks to this feature, the RTF algorithm offers superior classification accuracy compared to RF. The parameters of the machine learning models used in the study are given in Table 2.

Table 2. Parameters of machine learning algorithms used in the study.

Models	Parameters
SVM-Linear	penalty='l2' loss='squared_hinge' tol=0.0001 C=1.0 max_iter=1000
SVM-RBF	penalty='l2' loss='squared_hinge' C=1.0 Gamma=0.01
Random Forest	n_estimators=100 criterion='gini'
Rotation Forest	n_estimators=200 min_group=3 max_group=3 contract_max_n_estimators=500 remove_proportion=0.5

Performance Evaluation Metrics

Performance evaluation is an important aspect of the machine learning process. In this study, 4 different machine learning metrics are used for performance evaluation of the models. These metrics are expressed by the following mathematical formulae [47-48].

Accuracy (1): Represents the proportion of correct predictions among all predictions and indicates the overall model performance.

$$\mathrm{Accuracy} = (\mathrm{TN} + \mathrm{TP})/(\mathrm{TN} + \mathrm{FP} + \mathrm{TP} + \mathrm{FN}) \tag{1}$$

Precision (2): Measures how effectively a model identifies the positive class. However, using this measure alone may cause the model to minimise false positives.

$$\mathrm{Precision} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP}) \tag{2}$$

Recall (3): This measure shows how accurately a model predicts all positive instances in the dataset.

$$\mathrm{Recall} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \tag{3}$$

F1-score (4): It represents the harmonic mean between Precision and Recall and takes a value between 0 and 1. A value closer to 1 implies excellent Precision and Recall, while a value of 0 indicates that Precision and Recall are both 0. A high F1 score signifies a better balance between these two metrics.

$$\mathrm{F1} - \mathrm{score} = ((2 * \mathrm{Precision} * \mathrm{Recall}))/(\mathrm{Precision} + \mathrm{Recall}) \tag{4}$$

TP (True Positives): The number of positive observations that the model correctly predicts as positive.

FP (False Positive): The number of negative observations that the model incorrectly predicts as positive.

TN (True Negative): The number of negative observations that the model correctly predicts as negative.

FN (False Negative): The number of positive observations that the model incorrectly predicts as negative.

RESULTS

In the experimental assessment, the study explored the detection of *T. absoluta* damage by employing feature extraction through three distinct transfer learning techniques: DenseNet121, MobileNetV2, and EfficientNetB0. The dataset for *T. absoluta* damage used in these experiments was divided using the k-fold cross-validation methodology, with the k-fold value set to 10. This section details the findings of the research, emphasizing the analysis of images depicting *T. absoluta* damage. Deep features were extracted from these images using DenseNet121, MobileNetV2, and EfficientNetB0 architectures. Subsequently, these extracted deep features were utilized in the classification process through various machine learning algorithms, including Rotation Forest, Random Forest, Support Vector Machine-Linear (SVM-Linear), and Support Vector Machine-RBF (SVM-RBF). Tables 3, 4 and 5 provides a summary of the parameters employed in the Rotation Forest, Random Forest, SVM-Linear, and SVM-RBF models for the classification of images related to *T. absoluta* damage. Figure 4 shows the flowchart of the study.

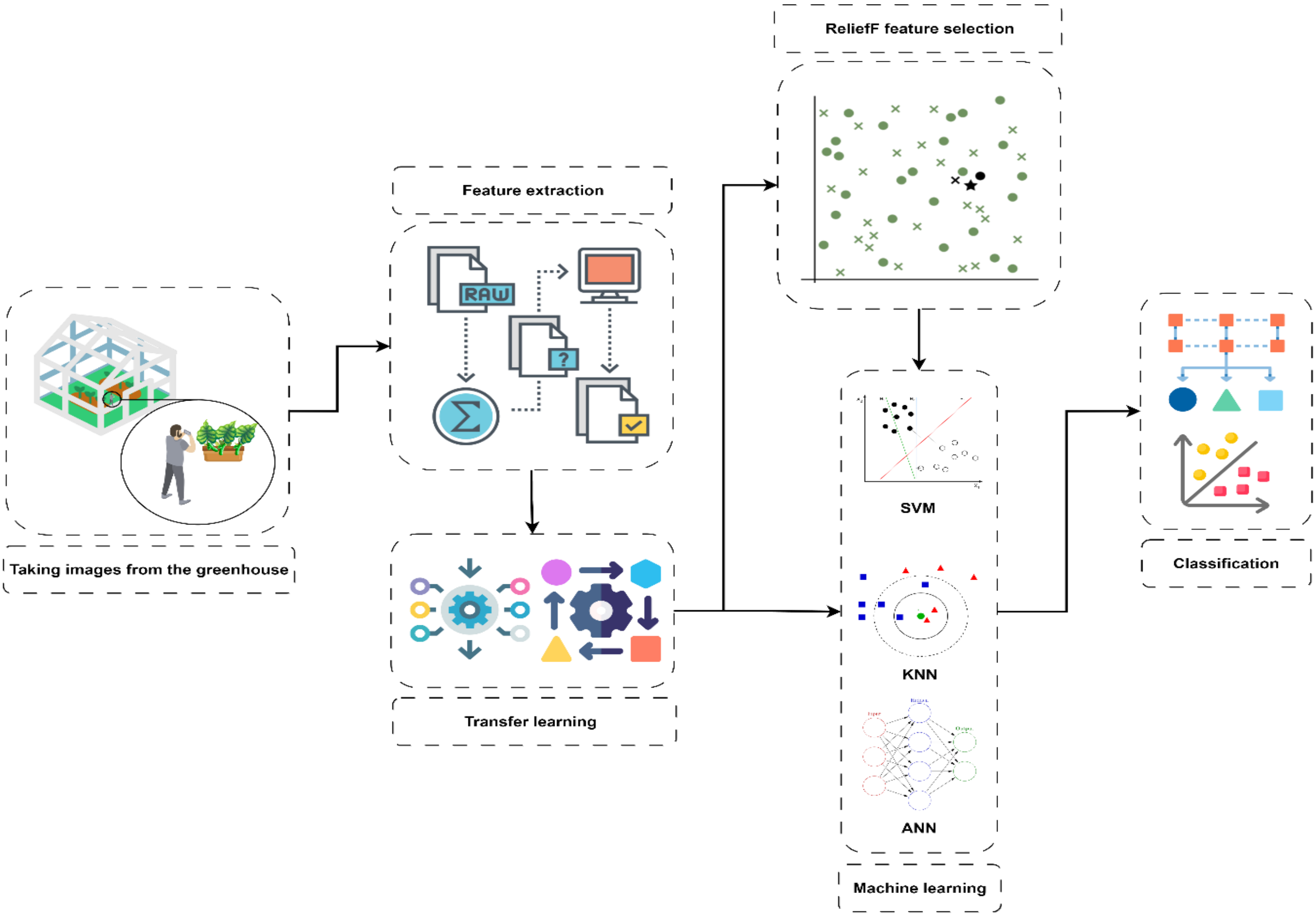


Figure 4. General flowchart of system.

Table 3. The performance results of Densenet121 deep features classification.

Models	Feature Selection	Class	Accuracy	Precision	Recall	F-Score
SVM-Linear	None	Healty	0.8756	0.6502	0.8400	0.7401
		Tomato leafminer	0.8756	0.9600	0.8800	0.9200
		Mean	0.8756	0.8051	0.8600	0.8305
SVM-RBF		Healty	0.8490	0.5960	0.7920	0.6874
		Tomato leafminer	0.8490	0.9412	0.8621	0.9214
		Mean	0.8490	0.7686	0.8270	0.8044
Random Forest		Healty	0.8514	0.5025	0.9012	0.6402
		Tomato leafminer	0.8514	0.9812	0.8423	0.9145
		Mean	0.8514	0.7418	0.8717	0.7773
Rotation Forest		Healty	0.8634	0.6145	0.8142	0.7021
		Tomato leafminer	0.8634	0.9510	0.8830	0.9123
		Mean	0.8634	0.7827	0.8486	0.8072
SVM-Linear	ReliefF	Healty	0.9251	0.7802	0.9311	0.8502
		Tomato leafminer	0.9251	0.9833	0.9212	0.9512
		Mean	0.9251	0.8817	0.9261	0.9007
SVM-RBF		Healty	0.9022	0.6812	0.9400	0.7902
		Tomato leafminer	0.9022	0.9801	0.8902	0.9401
		Mean	0.9022	09801	0.9151	0.8651
Random Forest		Healty	0.9444	0.8012	0.9901	0.8805
		Tomato leafminer	0.9444	1.0000	0.9301	0.9612
		Mean	0.9444	0.9006	0.9601	0.9208
Rotation Forest		Healty	0.9424	0.8500	0.9311	0.8912
		Tomato leafminer	0.9424	0.9802	0.9501	0.9613
		Mean	0.9424	0.9151	0.9406	0.9262

Table 3 presents a summary of the performance outcomes in classifying deep features extracted from DenseNet121 for the detection of *T. absoluta* damage. Various models, such as SVM-Linear, SVM-RBF, Random Forest, and Rotation Forest, were employed, all integrated with the ReliefF feature selection method. The evaluation metrics, encompassing accuracy, precision, recall, and F-score, are detailed for the two distinct classes: Healthy and *T. absoluta*. The table serves as a comprehensive reference, highlighting the effectiveness of each model-feature selection combination in discerning between the specified classes in the context of *T. absoluta* damage detection. For SVM-Linear and SVM-RBF, the 'None' feature selection method provides 87.56% and 84.90% accuracy respectively, with SVM-Linear achieving higher precision, recall and F-score. On the other hand, Random Forest and Rotation Forest show 85.14% and 86.34% accuracy respectively, with Rotation Forest showing superior precision, recall and F-score. The use of ReliefF feature selection improves the performance of SVM-Linear, resulting in an accuracy of 92.51% with higher precision, recall and F-score for both classes. Similar improvements are observed for ReliefF and SVM-RBF. Random Forest and Rotation Forest maintain their high accuracies of 94.44% and 94.24% respectively and provide higher precision, recall and F-score.

Figure 5 discusses the performance evaluation of various classification methods applied to the features extracted with DenseNet121 for the task of damage detection in *T. absoluta*. The paragraph summarizes different combinations of feature selection and classification techniques including ReliefF-SVM-Linear, ReliefF-SVM-RBF, ReliefF-Random Forest and ReliefF-Rotation Forest. Furthermore, Random Forest, SVM-Linear, SVM-RBF and Rotation Forest classification methods are evaluated separately without feature selection method. Throughout the paragraph, the important aspect of feature selection is emphasized, especially using the ReliefF method. Additionally, the classification methods Random Forest, SVM-Linear, SVM-RBF, and Rotation Forest are individually evaluated without the feature selection method. The crucial aspect of feature selection, particularly using the ReliefF method, is emphasized throughout the paragraph. The information suggests that employing feature selection improves the classification performance for SVM-Linear, SVM-RBF, Random Forest, and Rotation Forest models when applied to the extracted features from DenseNet121. The subsequent inclusion of a confusion matrix in Figure 5 aims to provide a detailed insight into the classification outcomes, enabling a more granular assessment of the models' effectiveness with and without feature selection.

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	171 20.65%	49 5.92%	220 77.73% 22.27%
omato Leafmine	13 1.57%	595 71.86%	608 97.86% 2.14%
SUM	184 92.93% 7.07%	644 92.39% 7.61%	766 / 828 92.51% 7.49%

ReliefF-SVM-Linear

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	149 18.00%	71 8.57%	220 67.73% 32.27%
omato Leafmine	10 1.21%	598 72.22%	608 98.36% 1.64%
SUM	159 93.71% 6.29%	669 89.39% 10.61%	747 / 828 90.22% 9.78%

ReliefF-SVM-RBF

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	176 21.26%	44 5.31%	220 80.00% 20.00%
omato Leafmine	2 0.24%	606 73.19%	608 99.67% 0.33%
SUM	178 98.88% 1.12%	650 93.23% 6.77%	782 / 828 94.44% 5.56%

ReliefF-Random Forest

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	186 22.46%	34 4.11%	220 84.55% 15.45%
omato Leafmine	14 1.69%	594 71.74%	608 97.70% 2.30%
SUM	200 93.00% 7.00%	628 94.59% 5.41%	780 / 828 94.20% 5.80%

ReliefF-Rotation Forest

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	109 13.16%	111 13.41%	220 49.55% 50.45%
omato Leafmine	12 1.45%	596 71.98%	608 98.03% 1.97%
SUM	121 90.08% 9.92%	707 84.30% 15.70%	705 / 828 85.14% 14.86%

Random Forest

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	144 17.39%	76 9.18%	220 65.45% 34.55%
omato Leafmine	27 3.26%	581 70.17%	608 95.56% 4.44%
SUM	171 84.21% 15.79%	657 88.43% 11.57%	725 / 828 87.56% 12.44%

SVM-Linear

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	130 15.70%	90 10.87%	220 59.09% 40.91%
omato Leafmine	35 4.23%	573 69.20%	608 94.24% 5.76%
SUM	165 78.79% 21.21%	663 86.43% 13.57%	703 / 828 84.90% 15.10%

SVM-RBF

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafmine	SUM
Healty	140 16.91%	80 9.66%	220 63.64% 36.36%
Tomato Leafmine	31 3.74%	577 69.69%	608 94.90% 5.10%
SUM	171 81.87% 18.13%	657 87.82% 12.18%	717 / 828 86.59% 13.41%

Rotation Forest

Figure 5. Confusion matrix of the classification obtained with and without the feature selection method of the features extracted with Densenet121.

Table 4. The performance results of MobileNetV2 deep features classification.

Models	Feature Selection	Class	Accuracy	Precision	Recall	F-Score
SVM-Linear	None	Healty	0.9239	0.8401	0.8712	0.8522
		Tomato leafminer	0.9239	0.9522	0.9411	0.9512
		Mean	0.9239	0.8961	0.9061	0.9017
SVM-RBF		Healty	0.8611	0.5741	0.8602	0.6802
		Tomato leafminer	0.8611	0.9720	0.8602	0.9201
		Mean	0.8611	0.7730	0.8602	0.8001
Random Forest		Healty	0.8575	0.5000	0.9401	0.6500
		Tomato leafminer	0.8575	0.9900	0.8401	0.9102
		Mean	0.8575	0.7450	0.8901	0.7801
Rotation Forest		Healty	0.8587	0.6000	0.8201	0.6903
		Tomato leafminer	0.8587	0.9502	0.8703	0.9108
		Mean	0.8587	0.7751	0.8452	0.8005
SVM-Linear	ReliefF	Healty	0.9795	0.9502	0.9700	0.9622
		Tomato leafminer	0.9795	0.9902	0.9800	0.9901
		Mean	0.9795	0.9702	0.9750	0.9761
SVM-RBF		Healty	0.9046	0.7202	0.9000	0.8041
		Tomato leafminer	0.9046	0.9703	0.9111	0.9432
		Mean	0.9046	0.8452	0.9055	0.8736
Random Forest		Healty	0.9408	0.8000	0.9800	0.8801
		Tomato leafminer	0.9408	0.9900	0.9300	0.9600
		Mean	0.9408	0.8950	0.9550	0.9200
Rotation Forest		Healty	0.9553	0.8710	0.9611	0.9111
		Tomato leafminer	0.9553	0.9901	0.9503	0.9711
		Mean	0.9553	0.9305	0.9557	0.9411

The information presented in Table 4 constitutes compelling evidence regarding the performance outcomes in classifying deep features extracted from MobileNetV2 in the context of *T. absoluta* damage detection. The table systematically details the results for diverse machine learning models, including SVM-Linear, SVM-RBF, Random Forest, and Rotation Forest, all integrated with the ReliefF feature selection method. Crucial performance metrics, encompassing accuracy, precision, recall, and F-Score, are meticulously documented for the two distinct classes: Healthy and *T. absoluta*. The noteworthy achievement of SVM-Linear with ReliefF is emphasized, showcasing an impressive accuracy of 97.95%, coupled with well-balanced precision, recall, and F-Score values for both classes. The varying degrees of performance observed across SVM-RBF, Random Forest, and Rotation Forest models, with ReliefF consistently enhancing classification outcomes, underscore the critical role of feature selection in the classification of *T. absoluta* damage using MobileNetV2 deep features.

Figure 6 discusses the performance evaluation of various classification methods applied to the features extracted with MobileNetV2 for the task of damage detection in *T. absoluta*. The paragraph summarises different combinations of feature selection and classification techniques including ReliefF-SVM-Linear, ReliefF-SVM-RBF, ReliefF-Random Forest and ReliefF-Rotation Forest. Furthermore, Random Forest, SVM-Linear, SVM-RBF and Rotation Forest classification methods are evaluated separately without feature selection method.

Training Set

TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	171 20.65%	49 5.92%	220 77.73% 22.27%
omato Leafmine	13 1.57%	595 71.86%	608 97.86% 2.14%
SUM	184 92.93% 7.07%	644 92.39% 7.61%	766 / 828 92.51% 7.49%

ReliefF-SVM-Linear

Training Set

TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	149 18.00%	71 8.57%	220 67.73% 32.27%
omato Leafmine	10 1.21%	598 72.22%	608 98.36% 1.64%
SUM	159 93.71% 6.29%	669 89.39% 10.61%	747 / 828 90.22% 9.78%

ReliefF-SVM-RBF

Training Set

TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	176 21.26%	44 5.31%	220 80.00% 20.00%
omato Leafmine	2 0.24%	606 73.19%	608 99.67% 0.33%
SUM	178 98.88% 1.12%	650 93.23% 6.77%	782 / 828 94.44% 5.56%

ReliefF-Random Forest

Training Set

TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	186 22.46%	34 4.11%	220 84.55% 15.45%
omato Leafmine	14 1.69%	594 71.74%	608 97.70% 2.30%
SUM	200 93.00% 7.00%	628 94.59% 5.41%	780 / 828 94.20% 5.80%

ReliefF-Rotation Forest

Training Set

TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	109 13.16%	111 13.41%	220 49.55% 50.45%
omato Leafmine	12 1.45%	596 71.98%	608 98.03% 1.97%
SUM	121 90.08% 9.92%	707 84.30% 15.70%	705 / 828 85.14% 14.86%

Random Forest

Training Set

TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	144 17.39%	76 9.18%	220 65.45% 34.55%
omato Leafmine	27 3.26%	581 70.17%	608 95.56% 4.44%
SUM	171 84.21% 15.79%	657 88.43% 11.57%	725 / 828 87.56% 12.44%

SVM-Linear

Training Set

TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	130 15.70%	90 10.87%	220 59.09% 40.91%
omato Leafmine	35 4.23%	573 69.20%	608 94.24% 5.76%
SUM	165 78.79% 21.21%	663 86.43% 13.57%	703 / 828 84.90% 15.10%

SVM-RBF

Training Set

TARGET / OUTPUT	Healty	Tomato Leafmine	SUM
Healty	140 16.91%	80 9.66%	220 63.64% 36.36%
Tomato Leafmine	31 3.74%	577 69.69%	608 94.90% 5.10%
SUM	171 81.87% 18.13%	657 87.82% 12.18%	717 / 828 86.59% 13.41%

Rotation Forest

Figure 6. Confusion matrix of the classification obtained with and without the feature selection method of the features extracted with MobileNetV2.

The information suggests that employing feature selection improves the classification performance for SVM-Linear, SVM-RBF, Random Forest, and Rotation Forest models when applied to the extracted features from MobileNetV2. The subsequent inclusion of a confusion matrix in Figure 6 aims to provide a detailed insight into the classification outcomes, enabling a more granular assessment of the models' effectiveness with and without feature selection.

Table 5. The performance results of EfficientNetB0 deep features classification.

Models	Feature Selection	Class	Accuracy	Precision	Recall	F-Score
SVM-Linear	None	Healty	0.9541	0.9000	0.9301	0.9111
		Tomato leafminer	0.9541	0.9714	0.9633	0.9732
		Mean	0.9541	0.9357	0.9467	0.9721
SVM-RBF		Healty	0.9019	0.7100	0.9000	0.7900
		Tomato leafminer	0.9019	0.9711	0.9002	0.9401
		Mean	0.9019	0.8405	0.9001	0.8650
Random Forest		Healty	0.9058	0.7500	0.8702	0.8101
		Tomato leafminer	0.9058	0.9601	0.9212	0.9412
		Mean	0.9058	0.8550	0.8957	0.8756
Rotation Forest		Healty	0.9106	0.7303	0.9111	0.8112
		Tomato leafminer	0.9106	0.9820	0.9103	0.9411
		Mean	0.9106	0.8561	0.9107	0.8761
SVM-Linear	ReliefF	Healty	0.9783	0.9400	0.9802	0.9601
		Tomato leafminer	0.9783	0.9900	0.9800	0.9900
		Mean	0.9783	0.9650	0.9801	0.9750
SVM-RBF		Healty	0.9070	0.7101	0.9200	0.8001
		Tomato leafminer	0.9070	0.9800	0.9000	0.9400
		Mean	0.9070	0.9450	0.9100	0.8700
Random Forest		Healty	0.9614	0.8901	0.9714	0.9222
		Tomato leafminer	0.9614	0.9901	0.9632	0.9700
		Mean	0.9614	0.9401	0.9673	0.9461
Rotation Forest		Healty	0.9589	0.9100	0.9300	0.9200
		Tomato leafminer	0.9589	0.9889	0.9789	0.9789
		Mean	0.9589	0.9494	0.9544	0.9494

Table 5 presents the performance results of classifying deep features extracted from EfficientNetB0 in the context of *T. absoluta* damage detection. Various machine learning models, such as SVM-Linear, SVM-RBF, Random Forest, and Rotation Forest, are evaluated in combination with the ReliefF feature selection method. The table comprehensively outlines essential performance metrics such as accuracy, precision, recall, and F-Score for the two distinct classes: Healthy and *T. absoluta*. SVM-Linear with ReliefF stands out with an impressive accuracy of 97.83%, demonstrating well-balanced precision, recall, and F-Score values for both classes. Across SVM-RBF, Random Forest, and Rotation Forest models, ReliefF consistently contributes to improved classification outcomes. The table serves as a valuable reference, providing insights into the comparative effectiveness of various models and feature selection methods in efficiently classifying damage caused by *T. absoluta* using EfficientNetB0 deep features.

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	171 20.65%	49 5.92%	220 77.73% 22.27%
omato Leafmine	13 1.57%	595 71.86%	608 97.86% 2.14%
SUM	184 92.93% 7.07%	644 92.39% 7.61%	766 / 828 92.51% 7.49%

ReliefF-SVM-Linear

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	149 18.00%	71 8.57%	220 67.73% 32.27%
omato Leafmine	10 1.21%	598 72.22%	608 98.36% 1.64%
SUM	159 93.71% 6.29%	669 89.39% 10.61%	747 / 828 90.22% 9.78%

ReliefF-SVM-RBF

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	176 21.26%	44 5.31%	220 80.00% 20.00%
omato Leafmine	2 0.24%	606 73.19%	608 99.67% 0.33%
SUM	178 98.88% 1.12%	650 93.23% 6.77%	782 / 828 94.44% 5.56%

ReliefF-Random Forest

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	186 22.46%	34 4.11%	220 84.55% 15.45%
omato Leafmine	14 1.69%	594 71.74%	608 97.70% 2.30%
SUM	200 93.00% 7.00%	628 94.59% 5.41%	780 / 828 94.20% 5.80%

ReliefF-Rotation Forest

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	109 13.16%	111 13.41%	220 49.55% 50.45%
omato Leafmine	12 1.45%	596 71.98%	608 98.03% 1.97%
SUM	121 90.08% 9.92%	707 84.30% 15.70%	705 / 828 85.14% 14.86%

Random Forest

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	144 17.39%	76 9.18%	220 65.45% 34.55%
omato Leafmine	27 3.26%	581 70.17%	608 95.56% 4.44%
SUM	171 84.21% 15.79%	657 88.43% 11.57%	725 / 828 87.56% 12.44%

SVM-Linear

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafminer	SUM
Healty	130 15.70%	90 10.87%	220 59.09% 40.91%
omato Leafmine	35 4.23%	573 69.20%	608 94.24% 5.76%
SUM	165 78.79% 21.21%	663 86.43% 13.57%	703 / 828 84.90% 15.10%

SVM-RBF

Training Set			
TARGET / OUTPUT	Healty	Tomato Leafmine	SUM
Healty	140 16.91%	80 9.66%	220 63.64% 36.36%
Tomato Leafmine	31 3.74%	577 69.69%	608 94.90% 5.10%
SUM	171 81.87% 18.13%	657 87.82% 12.18%	717 / 828 86.59% 13.41%

Rotation Forest

Figure 7. Confusion matrix of the classification obtained with and without the feature selection method of the features extracted with EfficientNetB0.

Figure 7 discusses the performance evaluation of various classification methods applied to the features extracted with EfficientNetB0 for the task of damage detection in *T. absoluta*. The paragraph succinctly outlines various combinations of feature selection and classification techniques, encompassing ReliefF-SVM-Linear, ReliefF-SVM-RBF, ReliefF-Random Forest, and ReliefF-Rotation Forest. Additionally, individual assessments are made for Random Forest, SVM-Linear, SVM-RBF, and Rotation Forest classification methods, each analyzed independently without the integration of a feature selection process. The presented information strongly implies that the incorporation of feature selection significantly enhances the classification efficacy across SVM-Linear, SVM-RBF, Random Forest, and Rotation Forest models when applied to features extracted from EfficientNetB0. SVM-Linear with ReliefF stands out with an impressive accuracy of 97.83%, demonstrating well-balanced precision, recall, and F-Score values for both classes. To provide a more detailed perspective, Figure 7 incorporates a confusion matrix, strategically included to offer an intricate understanding of the classification outcomes.

DISCUSSION

There are some studies in which machine learning algorithms have been used in the diagnosis of plant diseases by classifying the images through feature extraction. Table 6 is given to compare some of these studies with our study. Table 6 is given to compare some of these studies with our study. Table 6 contains the classification technique, data set, examined features and classification accuracy information of the listed studies. Accordingly, researchers have achieved high accuracy rates for different plant species using KNN, SVM, ANN, Bayes, Decision Tree, Random Forest, Multilayer Perceptron, Ada Boost, MLP, Logistic Regression, Linear Discriminant Analysis and CART machine learning algorithms.

This study was conducted to classify *T. absoluta* damage cases on tomato leaf images collected from private fields using mobile phones. In this study, feature extraction with deep learning and classification experiments with machine learning methods were performed. The dataset was divided using the k-fold cross-validation methodology, with the k-fold value set to 10. Deep features were extracted with transfer learning methods using DenseNet121, MobileNetV2 and EfficientNetB0 pre-trained models. After feature extraction, machine learning classification algorithms such as SVM-Linear, SVM-RBF, Random Forest and Rotation Forest were used to classify the deep color features. The results showed that the SVM-Linear algorithm performed the best among the others with an accuracy of 97.83%. Compared to the given studies, these results show that the use of color features in plant disease diagnosis is effective in achieving high accuracy rates. For example, KNN achieved an accuracy of 96.76% when classifying plant diseases in the Arkansas dataset, and Random Forest achieved 94% for tomato leaf images in a study using color histograms and local binary patterns. However, these studies often rely on standardized or public datasets, which lack the variability found in real-world conditions. The dataset used in this study, collected directly from the field, incorporates variability such as lighting conditions, natural damage patterns, and environmental factors, making it uniquely suited for practical applications. Furthermore, this study demonstrates the strength of combining deep learning-based feature extraction methods with machine learning classifiers, which outperformed hand-crafted feature approaches in the listed studies. Deep features extracted with DenseNet121, MobileNetV2, and EfficientNetB0 proved more effective at capturing subtle pest damage indicators, resulting in higher accuracy. This distinction underlines the novelty and practical value of this study in improving real-world pest detection methods. Future work in this area can contribute to more successful plant disease and pest damage classification by using larger data sets, testing different combinations of features, and developing more complex algorithms.

Table 6. Comparison literature studies and proposed method.

Study	Classification technique	Dataset	Features	Accuracy (%)
[49]	KNN	Arkansas plant disease data base and Reddit-plant leaf disease datasets	Color and texture (GLCM)	96.76
[50]	KNN	Groundnut leaf images	Color, texture, morphology	75.00

Cont. Table 6

[51]	ANN	Tomato and cotton leaf images	Contrast, correlation, energy, homogeneity, mean, standard deviation, variance	92.70
[52]	SVM	Rice plant	Color features extracted from each colorspace	94.65
[53]	Bayes Decision tree KNN SVM Random Forest	Maize leaf images	Shape, color and texture	77.56 77.46 76.16 74.35 79.23
[19]	Random Forest	Tomato leaf images	Color histograms, Hu moments, Haralick, local binary pattern features	94
[54]	Random Forest Bayes Multilayer perceptron AdaBoost SVM KNN	Cotton leaf images	Texture and color feature	92.56 84.29 96.69 90.08 97.52 91.73
[55]	MLP SVM	PlantVillage	Texture and color feature	97.78 97.23
[56]	ANN	Cabbage leaf images	Texture and color feature	92.00
[57]	Logistic Regression Linear Discriminant Analysis KNN CART Random Forest Bayes SVM	Paddy leaf images	Shape, color and texture	94.05 76.79 81.55 94.05 97.62 66.07 96.43
[58]	SVM KNN NN	Fusarium wilt disease in chickpea images	Texture and color feature	88.50 94.50 88.00
This Study	**SVM-Linear** **SVM-RBF** **Random Forest** **Rotation Forest**	**Tomato leaf images**	**Color features**	**97.83** **90.70** **96.14** **95.89**

Future Outlook, Pros, and Cons

Future studies should expand the dataset with images from diverse regions and growth stages to improve model generalizability. Exploring hybrid models that combine traditional feature extraction with deep learning could enhance accuracy, while lightweight, efficient models for smartphones or edge devices could increase accessibility for farmers. Incorporating multi-modal data, such as combining leaf images with environmental factors, may offer a more comprehensive pest damage prediction approach. This study’s strengths include the use of transfer learning techniques (DenseNet121, MobileNetV2, EfficientNetB0) to extract discriminative features, achieving a high accuracy of 97.83%. Transfer learning also reduces the need for extensive labeled data and accelerates training. The method is adaptable to different plants, pests, and diseases, making it scalable for agricultural pest management. However, the computational demands of transfer learning may limit its use in resource-constrained settings, and creating a custom dataset requires significant effort in data collection and labeling. While effective for *T. absoluta*, further validation is needed for its application to other pests and crops.

CONCLUSION

Studies using deep learning techniques to automatically detect plant diseases and pests have increased in recent years. Various deep learning algorithms have been used in previous studies. In this study, unlike previous studies, a new data set was originally created for the purpose of *T. absoluta* damage detection, deep features were extracted from these images using three different transfer learning techniques (DenseNet121, MobileNetV2 and EfficientNetB0) and then, for the purpose of classification of these extracted deep features, various machine learning algorithms (Rotation Forest, Random Forest, Support Vector Machine-Linear (SVM-Linear) and Support Vector Machine-RBF (SVM-RBF) were used. The results obtained showed that the method we proposed for determining *T. absoluta* damage achieved high performance with an accuracy rate of 97.83% in detecting *T. absoluta* damage. The proposed method is built as a combination of feature extraction with CNN transfer learning and traditional machine learning approaches, and is a newly used method in the classification of plant pests and diseases. This method is successful in capturing complex spatial relationships and patterns in images of tomato leaves, allowing subtle pest damage indicators to be distinguished. This innovation is of great importance as it increases the accuracy and efficiency of disease detection compared to traditional methods. Therefore, the successful identification of *T. absoluta* damage suggests that this method can be used to detect various pests and diseases in other plant species. The significance of this study extends beyond accurate pest detection; it highlights the feasibility of implementing cost-effective, real-world solutions for agricultural pest management. By demonstrating the ability to use readily available mobile devices for data collection, the study provides a practical framework that can be adopted in resource-limited environments. In the proposed method, a fast, reliable and robust classification of *T. absoluta* damages was achieved. In this way, the classification, which is an intelligent and automatic system, is made with easily available equipment without the use of costly and special imaging devices. Furthermore, this approach can be integrated into broader precision agriculture systems, paving the way for automated monitoring and targeted pest control applications. For instance, combining this method with variable rate spraying systems could minimize pesticide use, reduce environmental impact, and optimize crop yields. Additionally, the proposed method can assist experts or farmers in their decision-making processes. Thus, it can help eliminate problems such as high labor costs and time losses. By adding autonomous and variable rate spraying systems to the system, it will be possible to spray only pest damaged or diseased areas. In this case, diagnosis of plant pests and diseases, determination of the severity and course of infestation/infection will be faster and more practical, and more effective plant protection will be possible with early intervention.

Funding: This research received no external funding.
Acknowledgments: None
Conflicts of Interest: The authors declare no conflict of interest

REFERENCES

1. Özgüven MM. Digital Agriculture Technologies Used in Horticultural Cultivation. J Agric Mach Sci. 2023; 19(3):174-93.
2. Ozguven MM. The digital age in agriculture. CRC Press Taylor & Francis Group LLC; 2023. ISBN 978-103-23-8577-8.
3. Savary S, Willocquet L. Simulation modeling in botanical epidemiology and crop loss analysis. Plant Health Instr. 2014. 173 p.
4. Avelino J, Cristancho M, Georgiou S, Imbach P, Aguilar L, Bornemann G. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. Food Secur. 2015;7(2):303–21.
5. Kang J, Zhao L, Wang K, Zhang K. Research on an improved YOLOv8 image segmentation model for crop pest. Adv Comput Signals and Syst. 2023;7(3). DOI: 10.23977/acss.2023.070301. ISSN 2371-8838.
6. Desneux N, Luna MG, Guillemaud T, Urbaneja A. The invasive South American tomato pinworm, Tuta absoluta, continues to spread in Afro-Eurasia and beyond: the new threat to tomato world production. J Pest Sci. 2011;84:403-8.
7. Caparros Megido R, Haubruge E, Verheggen FJ. First evidence of Deuterotokous Parthenogenesis in the tomato leafminer, Tuta absoluta (Meyrick) (Lepidoptera: Gelechiidae). J Pest Sci. 2012;85:409–12. DOI: 10.1007/s10340-012-0458-6.
8. Terzidis AN, Wilcockson S, Leifert C. The tomato leaf miner (Tuta absoluta): Conventional pest problem, organic management solutions? Org Agr. 2014;4:43–61. https://doi.org/10.1007/s13165-014-0064-4.
9. Zekeya N, Ndakidemi PA, Chacha M, Mbega E. Tomato leafminer, Tuta absoluta (meyrick 1917), an emerging agricultural pest in sub-saharan Africa: current and prospective management strategies. Afr J Agric Res. 2017;12(6):389-96.
10. Rubanga DP, Loyani LK, Richard M, Shimada S. A deep learning approach for determining effects of Tuta Absoluta in tomato plants. arXiv preprint. 2020. https://doi.org/10.48550/arXiv.2004.04023.
11. Uygun T, Ozguven MM, Yanar D. A new approach to monitor and assess the damage caused by two-spotted spider mite. Exp Appl Acarol. 2020;82(3):335-46. Doi: 10.1007/s10493-020-00561-8.
12. Ozguven MM, Altas Z. A new approach to detect mildew disease on cucumber (Pseudoperonospora cubensis) Leaves with Image Processing. J Plant Pathol. 2022. https://doi.org/10.1007/s42161-022-01178-z.

13. Ozguven MM, Yanar Y. The technology uses in the determination of sugar beet diseases. In: Misra V, Srivastava S, Mall AK, editors. Sugar Beet Cultivation, Management and Processing. Springer, Singapore; 2022. Doi: 10.1007/978-981-19-2730-0_30.
14. Dönmez E, Kılıçarslan S, Közkurt C, Diker AB, Demir FB, Elen A. Identification of haploid and diploid maize seeds using hybrid transformer model. Multimed Syst. 2023;29(6):3833–45. DOI: 10.1007/s00530-023-01174-y.
15. Adem K, Közkurt C. Defect detection of seals in multilayer aseptic packages using deep learning. Turk J Electr Eng Comput Sci. 2019;27(6):4220–30. Doi: 10.3906/elk-1903-112.
16. Kiliçarslan S, Közkurt C, Baş S, Elen A. Detection and classification of pneumonia using novel Superior Exponential (SupEx) activation function in convolutional neural networks. Expert Syst with Appl. 2023;217:119503.
17. Közkurt C, Kiliçarslan S, Baş S, Elen A. α-SechSig and α-TanhSig: Two novel non-monotonic activation functions. Soft Comput., 2023;27(24), 18451–67. https://doi.org/10.1007/s00500-023-09279-2.
18. Ahmed SF, Alam MSB, Hassan M, et al. Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artif Intell Rev 2023;56, 13521–617. https://doi.org/10.1007/s10462-023-10466-8.
19. Basavaiah J, Arlene Anthony A. Tomato leaf disease classification using multiple feature extraction techniques. Wirel. Pers. Commun., 2020;115, 633-51. https://doi.org/10.1007/s11277-020-07590-x.
20. Şahin YS, Erdinç A, Bütüner AK, Erdoğan H. Detection of Tuta absoluta larvae and their damages in tomatoes with deep learning-based algorithm. Int. J. Comput. Sci. Res., 2023. https://doi.org/10.47164/ijngc.v14i3.1287.
21. Georgantopoulos PS, Papadimitriou D, Constantinopoulos C, Manios T, Daliakopoulos IN, Kosmopoulos D. A Multispectral dataset for the detection of Tuta Absoluta and Leveillula Taurica in tomato plants. Smart Agric. Technol., 2023;4. 100146. https://doi.org/10.1016/j.atech.2022.100146.
22. Mkonyi L, Rubanga D, Richard M, Zekeya N, Sawahiko S, Maiseli B, et al. Early identification of Tuta absoluta in tomato plants using deep learning. Sci. Afr., 2020;10, e00590. https://doi.org/10.1016/j.sciaf.2020.e00590.
23. Loyani, LK, Bradshaw, K, Machuve, D. Segmentation of Tuta Absoluta's damage on tomato plants: a computer vision approach. Appl. Intell., 2021;35(14), 1107–27. https://doi.org/10.1080/08839514.2021.1972254.
24. Giakoumoglou N, Pechlivani E-M, Frangakis N, Tzovaras D. Enhancing Tuta absoluta detection on tomato plants: ensemble techniques and deep learning. AI, 2023;4(4), Article 4. https://doi.org/10.3390/ai4040050.
25. Diker A, Elen A, Közkurt C, Kılıçarslan S, Dönmez E, Arslan K, et al. An effective feature extraction method for olive peacock eye leaf disease classification. Eur. Food Res. Technol., 2023. https://doi.org/10.1007/s00217-023-04386-8.
26. Abdel-Razek AS, Masry SHD, Sadek HE, Gaafar HE. Efficacy of Trichogramma wasps for controlling tomato leaf miner Tuta Absoluta, Arch. Phytopathol. Plant Prot., 2019;52:5-6, 443-57. DOI: 10.1080/03235408.2019.1634782.
27. Pandey M, Bhattarai N, Pandey P, Chaudhary P, Katuwal DR, Khanal D.A Review on biology and possible management strategies of tomato leaf miner, *Tuta Absoluta* (Meyrick), Lepidoptera: Gelechiidae in Nepali. 2023. Available at SSRN: https://ssrn.com/abstract=4525117 or http://dx.doi.org/10.2139/ssrn.4525117.
28. Kılıç T. Plant Pest Agricultural Struggle Technical Instructions. Department of Plant Health Research, 28-29. General Directorate of Agricultural Research and Policies. 2008.
29. Desneux N, Wajnberg E, Wyckhuys KAG, Burgio G, Arpaia S, Narváez-Vasquez CA, et al. Biological invasion of European tomato crops by Tuta absoluta: ecology, geographic expansion and prospects for biological control. J Pest Sci 2010;83, 197–215. https://doi.org/10.1007/s10340-010-0321-6.
30. Kılıçarslan S, Kılıçarslan S. A comparative study of bread wheat varieties identification on feature extraction, feature selection and machine learning algorithms. Eur. Food Res. Technol., 2023;250(1):135–149. https://doi.org/10.1007/s00217-023-04372-0
31. Ozer I, Kocak I, Cetin O, Karaca AC, Ozer CK, Gorur K. Towards investigation of transfer learning framework for Globotruncanita genus and Globotruncana genus microfossils in Genus-Level and Species-Level prediction. Eng. Sci. Technol., 2023;48, 101589. https://doi.org/10.1016/j.jestch.2023.101589.
32. Torrey L, Shavlik JW. Transfer Learning. IGI Global eBooks; 2010. doi:10.4018/978-1-60566-766-9.ch011.
33. Huang G, Liu S, Maaten L, Weinberger KQ. CondenseNet: An efficient DenseNet using learned group convolutions. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018. doi:10.1109/cvpr.2018.00291.
34. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications arXiv.org; 2017. URL https://arxiv.org/abs/1704.04861.
35. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. MobileNetV2: Inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018. doi:10.1109/cvpr.2018.00474.
36. Tan M, Le QV. EfficientNetV2: Smaller models and faster training arXiv.org; 2021. URL https://arxiv.org/abs/2104.00298. DOI: 10.48550/arXiv.2104.00298.
37. Osuna E, Freund R, Girosit F. Training support vector machines: an application to face detection. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition;1997. doi:10.1109/cvpr.1997.609310.
38. MohammedAbuBasim N, Sathyabalan P, Suresh P. Analysis of EEG Signals and facial expressions to detect drowsiness and fatigue using gabor filters and SVM linear classifier. J. Comput. Appl., 2015;115 11, 9–14. doi:10.5120/20194-2433.

39. Elen A, Baş S, Közkurt C. An adaptive gaussian kernel for support vector machine. Arab. J. Sci. Eng., 2022;47(8), 10579–88. https://doi.org/10.1007/s13369-022-06654-3.
40. Altun H, Polat G. On the comparison of classifiers' performance in emotion classification: Critiques and suggestions. In: 2008 IEEE 16th Signal Processing, Communication and Applications Conference; 2008. doi:10.1109/siu.2008.4632592.
41. Kecman V. Learning and soft computing. Cambridge, MA: The MIT Press; 2001.
42. Caie PD, Dimitriou N, Arandjelović O. Chapter 8—Precision medicine in digital pathology via image analysis and machine learning. In: Cohen S (ed) Artificial intelligence and deep learning in pathology. Elsevier; 2021. pp 149–173. https:// doi. org/ 10. 1016/ B978- 0- 323- 67538-3. 00008-7.
43. Özlem A, Güngör O. Classification of multispectral images using random forest algorithm. Hkmojjd., 2012;106:139–46.
44. Khamar M, Eftekhari M. Multi-manifold based rotation forest for classification. Appl. Soft Comput., 2018; 68:626–35. https://doi.org/10.1016/j.asoc.2018.04.026.
45. Liu KH, Huang DS. Cancer classification using Rotation Forest. Comput. Biol. Med., 2008;38(5):601–10. https://doi.org/10.1016/j.compbiomed.2008.02.007.
46. Rodriguez J, Kuncheva L, Alonso C. Rotation Forest: A new classifier Ensemble Method. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006;28(10):1619–30. https://doi.org/10.1109/tpami.2006.211.
47. Japkowicz N, Shah M. Performance evaluation in machine learning. In: El Naqa, I., Li, R., Murphy, M. (eds) Machine Learning in Radiation Oncology. Springer, Cham. 2015;41–56. https://doi.org/10.1007/978-3-319-18305-3_4.
48. Restrepo-Arias JF, Branch-Bedoya JW, Awad G. Plant disease detection strategy based on image texture and bayesian optimization with small neural networks. Agriculture. 2022;12, 1964. https://doi.org/10.3390/agriculture12111964.
49. Hossain E, Hossain MF, Rahaman MA. A Color and texture based approach for the detection and classification of plant leaf disease using KNN classifier. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE); 2019. doi:10.1109/ecace.2019.8679247.
50. Vaishnnave MP, Devi KS, Srinivasan P, Arutperumjothi G. Detection and classification of groundnut leaf diseases using KNN classifier. 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN); 2019;1-5. doi: 10.1109/ICSCAN.2019.8878733.
51. Kumari CU, Prasad SJ, Mounika G. Leaf Disease Detection: Feature extraction with K-means clustering and classification with ANN. 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC); 2019. doi:10.1109/iccmc.2019.8819750.
52. Shrivastava VK, Pradhan MK. Rice plant disease classification using color features: a machine learning paradigm. J. Plant Pathol., 2020;103:17–26. doi:10.1007/s42161-020-00683-3.
53. Panigrahi KP, Das H, Sahoo AK, Moharana SC. Maize leaf disease detection and classification using machine learning algorithms. Adv. Intell. Syst. Comput., 2020;659–69. doi:10.1007/978-981-15-2414-1_66.
54. Patil BM, Burkpalli V. A perspective view of cotton leaf image classification using machine learning algorithms using WEKA. Adv. Hum. -Comput., 2021;1–15. https://doi.org/10.1155/2021/9367778.
55. Altalak M, Uddin MA, Alajmi A, Rizg A. A hybrid approach for the detection and classification of tomato leaf diseases. Appl. Sci., 2022;12(16):8182. doi:10.3390/app12168182.
56. Myna AN, Manasvi K, Pavan JK, Rakshith HS, Yuktha DJ. Classification and detection of cabbage leaf diseases from images using deep learning methods. Autom Control Intell Syst., 2023;11(1): 1-7. doi: 10.11648/j.acis.20231101.11.
57. Saminathan K, Sowmiya B, Chithra DM. Multiclass classification of paddy leaf diseases using Random Forest classifier. J. Image Graph., 2023;11(2):195–203. doi:10.18178/joig.11.2.195-203.
58. Hayit T, Endes A, Hayit F. KNN-based approach for the classification of fusarium wilt disease in chickpea based on color and texture features. Eur. J. Plant Pathol., 2023. https://doi.org/10.1007/s10658-023-02791-z.