SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 8, No. 4, 592-605, 2025
DOI: 10.35377/saucis...1626178

RESEARCH ARTICLE

Enchancing Apple Plant Leaf Disease Detection Performance with Transfer Learning Methods

Alican Doğan[1,*], Cemal Yüksel [2,]

[1,2] Department of Management Information Systems, Bandirma Onyedi Eylul University, Balıkesir, Türkiye, ror.org/ 02mtr7g38

Corresponding author:
Alican Doğan,
Department of Management Information Systems,
Bandirma Onyedi Eylul University,
Balıkesir, Türkiye
alicandogan@bandirma.edu.tr

Article History:
Received: 24.01.2025
Revised: 10.08.2025
Accepted: 01.09.2025
Published Online: 13.10.2025

ABSTRACT
It is very important in agriculture to detect plant diseases and develop recovery solutions to produce more crops and improve efficiency. Enhancements in automated disease detection and analysis can offer significant advantages for taking prompt action, enabling interventions at earlier stages to treat the disease and prevent its spread. This proactive approach could help minimize damage to crop yields. This research aims to improve classification performance for apple leaf disease detection using transfer learning approaches. It discriminates sick apple plants from healthy counterparts by implementing image processing with apple leaf photographs. This study applies traditional machine learning methods for the apple plant disease detection task, and the classification achievement scores are maximized with transfer learning techniques. The experiments are conducted on a real-world data set including 3164 apple leaf images. Five classical machine learning classifiers (Support Vector Machine, Decision Tree, Random Forest, Logistic Regression, and Naïve Bayes) were compared against five deep transfer learning architectures (ResNet50, EfficientNetB0, InceptionV3, DenseNet121, and Xception) under stratified 10-fold cross-validation. As a result, those experiments reveal that transfer learning methods, especially EfficientNetB0, have significantly improved classification accuracy for this task.

Keywords: Machine learning, Transfer learning, Plant disease detection, Agriculture

1. Introduction

The condition of a plant can often be assessed by observing the appearance of its leaves [1]. Different harmful infections can be observed over time as the leaf's natural shape becomes distorted, displaying unusual and abnormal structures [2]. Apple Leaf Disease (ALD) can contribute to yield losses, decrease the quality of apples, and reduce the efficiency of the photosynthesis process. The primary and common symptoms of ALD, such as wilting, discoloration, lesions, and spotting, can be quickly identified through visual inspection [3]. Visual inspection poses a challenge in vast apple orchards because of the extensive area that must be surveyed. This constraint can be mitigated by employing drones outfitted with cameras and visual sensors, facilitating effective image acquisition and surveillance across the orchard. Alternative methods, including field surveys [4], image-sharing websites, and pre-existing datasets [5], can be employed to gather photographs of apple leaves for analysis. Nonetheless, other issues emerge after the collection of these photos. Factors such as insufficient lighting, inconsistent illumination, irregular leaf orientation, overlap, and impediments during the capture process hinder the proper diagnosis of ALD [6].

Agricultural disease diagnosis on large plantations is still a manpower- and expertise-intensive operation and frequently limited by the availability and expertise of people for inspections [7]. A range of intelligent diagnosis systems has been introduced in response, with encouraging results in numerous controlled conditions [8]. A few of these systems are even meant to assist the farmers through mobile interfaces without requiring agronomic-support level expertise or expert-level training [9], [10].

Of all the above methods, the predictive model methods—particularly the artificial intelligence-oriented methods—have been quite accurate in detecting diseases such as ALD [11]. However, despite their success, these models are often computationally intensive, which limits their scalability and real-time field deployment [12]. These approaches can forecast the disease with high precision due to the use of advanced predictive algorithms. However, a key limitation is their intensive computational demand, which limits their practicality in real-world deployment [13]-[14]. This study analyzes the existing methods' strengths and weaknesses, identifying opportunities to develop more computationally efficient and cost-effective solutions suitable for broader applications.

The main contributions of this article are as follows: i) the study visualizes class distributions, uses alluvial diagrams to detect potential class overlaps. ii) It utilizes transfer learning models like ResNet, Inception, EfficientNet, DenseNet, and Exception

Cite as: A. Doğan and C. Yüksel, "Enchancing apple plant leaf disease detection performance with transfer learning methods" *Sakarya University Journal of Computer and Information Sciences,* vol. 8, no. 4, pp. 592-605, 2025. doi: 10.35377/saucis...1626178

to classify leaves affected by scab, rust, multiple diseases, and healthy conditions. They apply stratified 10-fold cross-validation to address class imbalance in the dataset, leading to improved model performance. The model achieved notable precision and overall accuracy (99.93%).
The organization of the article is planned as follows. Section 2 mentions related research studies about apple plant leaf disease detection. Section 3 provides information about the conventional machine learning (ML) classification methods and transfer learning concepts. In addition, it provides detailed information about the EfficientNet method. The results of experimental studies are presented in Section 4. This section demonstrates the comparison among distinct classification techniques. Eventually, Section 5 concludes the research study and proposes possible future studies.

2. Related Work

With the growing threat of plant diseases to global food security, early and accurate detection has become a research priority in precision agriculture [15], [16]. Recent advances in artificial intelligence have enabled the development of automated systems that analyze large-scale leaf image datasets to diagnose plant health accurately [17], [18]. Such AI-based solutions support early intervention and disease management and contribute to long-term sustainability in agricultural practices by reducing unnecessary crop losses [19].

The leaf is a primary indicator of an apple plant's health [20]. With advances in various image-capturing technologies, obtaining leaf images is now relatively straightforward [21]. Analyzing these images can provide valuable insights into factors such as tree health, fruit quality, potential yield reduction, disease spread, management costs, and the overall environmental impact of diseases like ALD. Diseases, including circular spots and twisted leaf shapes, are early signs of health deterioration [22]. Advanced indicators of ALD include scabs and lesions, visible in various forms, which help diagnose the disease and facilitate early intervention for effective disease management [23].

Major advances in ML technology have enabled numerous architectures and capabilities for automated systems to rapidly and accurately detect plant diseases [24]. These systems can use computer learning to analyze and interpret plant photographs to facilitate early detection, disease management, and reduce crop loss. In [25], a study combining deep learning (DL) and machine learning tools was conducted to detect and demonstrate ALD. The study uses a predictive representation model based on a traditional Kaggle dataset. The study omits the procedural steps of image acquisition, preprocessing, and feature extraction required to enrich the input dataset. Instead, it focuses on model training and refinement, explaining ALD detection using machine learning and deep learning methods.

Experts have come up with several ways to find plant diseases that use image processing, machine learning, and machine learning [16]-[26]-[27]. Image processing methods make it easier to get important visual features from leaf images, like texture, color, and shape, which show that the plant is sick [28–29]. Then, machine learning and machine learning models look at these features to figure out what kind of plant disease they are or to guess what kind of plant disease they are. Experts can better control disease in agriculture by using these methods to determine what's wrong with plants.

Studies on transfer learning in detecting apple leaf disease have achieved high accuracy rates using different deep learning architectures and popular datasets such as Kaggle and PlantVillage. For example, an Inception-based model achieved a validation accuracy of 99.03% on the Kaggle dataset [30]. In a study comparing eight different pre-trained CNN architectures, MobileNet stood out with the highest accuracy value of 97%, despite its lighter structure [31]. [32] The ResNet50v2-based TransferNet model was adapted to real-world challenges under varying lighting conditions, reporting 91.63% accuracy and 91.03% F1 score. A report indicated that AlexNet achieved 99.56% accuracy on the PlantVillage dataset [33]; similarly, another study combining the EfficientNetV2S architecture with data augmentation techniques achieved a success rate of 99.21% [34]. These models provide important application insights regarding generalizability and real-time deployment while balancing model complexity and computational efficiency [35]. These findings highlight that transfer learning has strong potential for apple leaf disease detection regarding accuracy and efficiency; however, further optimization is needed regarding architectural complexity–performance balance and real-time applications.

Recent studies (such as [36-38]) have shown the use of DL algorithms to improve the predictive accuracy in identifying plant diseases. Due to their ability to learn complex patterns and features in large datasets, DL algorithms have outperformed traditional machine learning methods. In addition, segmentation techniques for ALD detection aim to separate the foreground (e.g., infected leaf areas) from the background more effectively. For example, [39] addressed segmentation, introducing semantic segmentation to facilitate a comprehensive analysis of infected areas and accurately distinguish and extract lesion features. Their work collected data from different spatial scales using a network model with a spatial pyramid pooling layer. This layer recognizes the model's complex features and improves the segmentation's accuracy. This method, leveraging transfer learning, successfully improved ALD detection, highlighting the effectiveness of advanced segmentation. This method, leveraging transfer learning, successfully improved ALD detection, highlighting the importance of segmentation in improving ALD detection. Generalizability issues and class imbalance are persistent problems in the literature. DL models work well for controlled datasets but may have problems when applied to other species or environmental conditions [40].

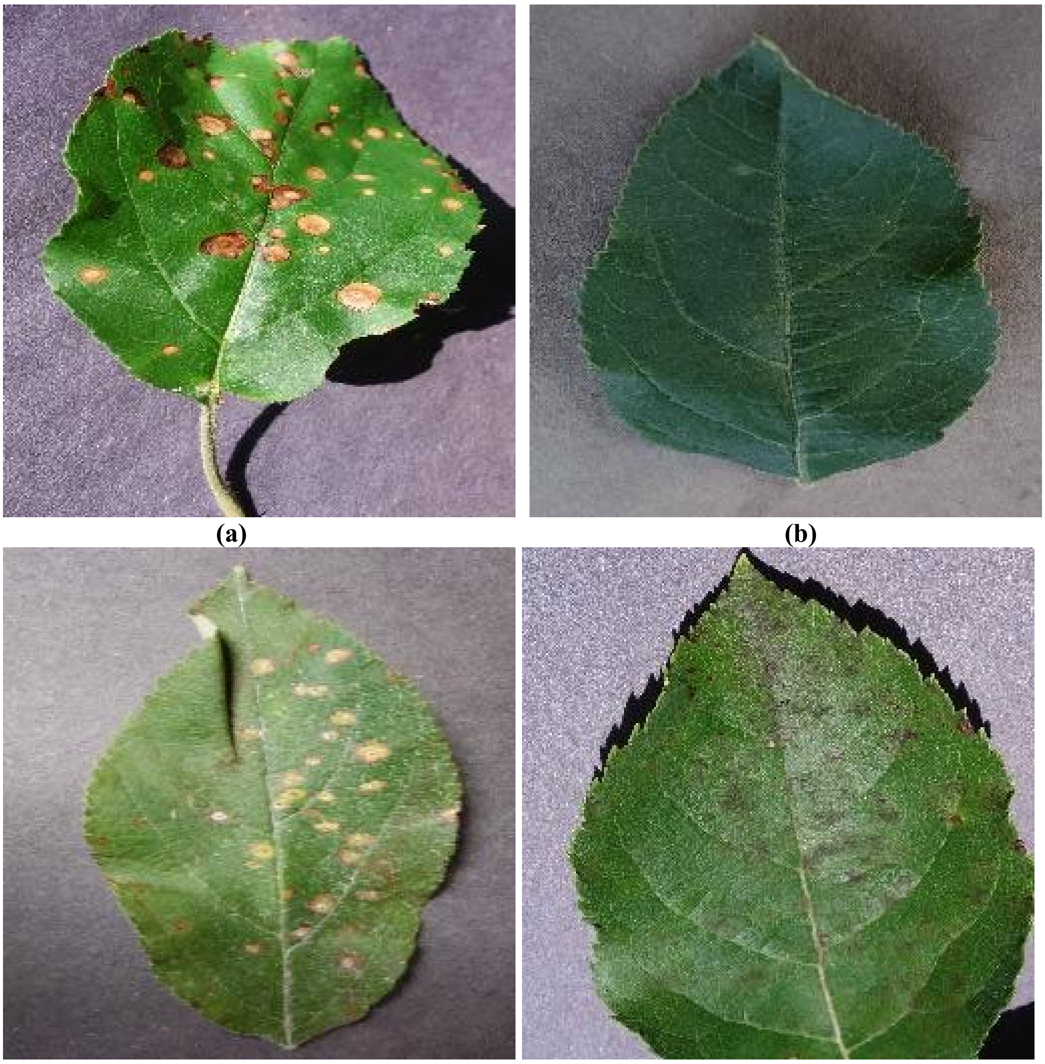
ML and DL algorithms for classifying plant leaf diseases cover a variety of crops and locations. Several studies have been applied to various plant species, including tomato, apple [41], rice [42], pepper [43], lemon [44], olive [45], cherry [46], apricot [47], strawberry, and peach [48]. This study prioritizes identifying ALDs, given the widespread cultivation and economic importance of apples globally, the scale of crop production, and its important role in international agriculture.

3. Materials and Methods

3.1 Dataset Description

Battal's Leaf Disease Images dataset, created in 2024, is a valuable resource for plant pathology and agricultural research. Containing over 7,000 photographs of healthy and diseased leaves from two plant species, this dataset can be used to train machine learning models to identify plant diseases. This dataset is an important tool for developing and testing artificial intelligence algorithms. It provides a large, standardized collection of annotated plant images that allows researchers to build and validate effective predictive models for various plant diseases and conditions. The dataset comprises labeled photographs of apples and grapes and is highly diverse. This dataset allows researchers to train models on diseased and healthy plants to improve classification accuracy for disease detection. This cross-species label diversity is useful in computer vision, machine learning, and other applications, enabling the development of robust and adaptable plant disease diagnostic systems.

In this research, only the images of apple leaves are used from this dataset. It includes 3164 images of apple leaves with four different class labels: healthy, black rot, rust, and scab. Black rot, rust, and scab categories represent diseases, while healthy images express the absence of disease in apple leaves. Figure 1 shows sample apple leaf images from those different groups. The input data consists of 256×256×3 RGB images labeled into four classes (healthy, black rot, rust, scab). Leaf images labeled as black rot, rust, and scab are marked as diseased, which makes the data suitable for a binary classification problem. For classical ML models, each image is flattened into a 196,608-dimensional feature vector (256 × 256 × 3). Transfer learning architectures receive the normalized RGB images directly as input, without manual feature extraction.



(a) **(b)**

(b) **(d)**

Figure 1. Sample images of apple leaves with different categories (a) for black rot, (b) for healthy, (c) for rust, and (d) for scab.

A preprocessing step involving resizing and rescaling operations is crucial to classify images effectively. These steps maximize the algorithm's ability to capture key characteristics of plant diseases while minimizing errors.

Resizing: Input images are resized to 256x256 pixels for uniformity and efficiency. This standardization simplifies computational overhead and aligns the dataset so the model can process it consistently.

Rescaling: After resizing, rescaling is applied by dividing the value of each pixel by 255. This process normalizes the pixel values, converting them from a range of 0-255 to a range of 0-1. Normalization preserves the distribution of values, ensuring the preservation of key features in the image and enhancing the model's ability to learn effectively.

This preprocessing step improves the overall accuracy and reliability of the classification task by preparing images as input for ML or DL models. Information regarding the overall data distribution by category is provided in Table 1, Table 2, Figure 2, and Figure 3. The numbers indicate the number of samples from each class label.

Table 1. Data Distribution for All Classes in Apple Plant Leaf Dataset

Classes of Data	**# of instances**	**%**
Apple_Black_rot	620	19.2
Apple_healthy	1640	50.7
Apple_rust	275	8.5
Apple_scab	629	19.5
Total	3164	100

Table 2. Data Distribution for All Classes in Apple Plant Leaf Dataset

Classes of Data	**# of instances**	**%**
Healthy Apple	1640	50,7
Sick Apple	1524	49,3
Total	3164	100

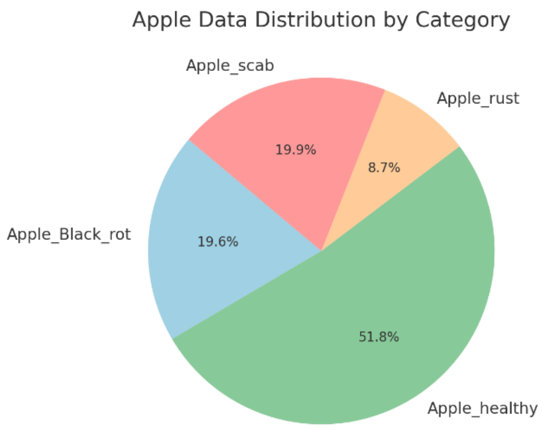


Figure 2. Pie Chart Showing Apple Data Distribution with All Classes

3.2 Transfer Learning

Artificial Neural Networks (ANNs) are loosely inspired by the structure and functioning of neural networks in the mammalian brain [49]. These networks consist of multiple layers of interconnected nodes, or neurons, that utilize activation functions to process information [50]. Data enters the system through the input layer, which transmits to subsequent layers. The layers transform the data using weighted connections, enabling the network to learn patterns and relationships. Finally, the processed information is delivered to the output layer, providing the desired results.

Convolutional Neural Networks (CNNs) are a specialized feed-forward neural network designed to process structured data, such as images, by leveraging convolutional operations to extract meaningful features [51]. The development of CNNs owes much to LeNet-5, a seven-layer convolutional network incorporating backpropagation and adaptive weight mechanisms to optimize parameters. Most modern CNN architectures build upon the foundational principles established by LeNet-5.

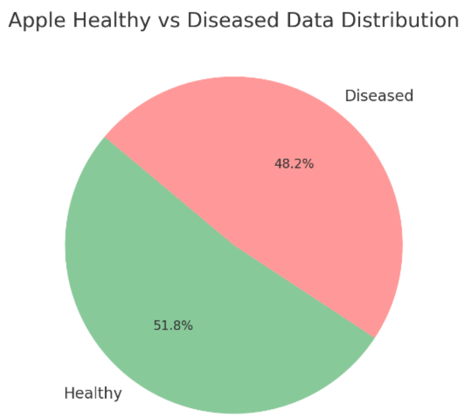


Figure 3. Pie Chart Showing Apple Data Distribution with Only Healthy and Diseased Leaves

The key advantages of CNNs can be described as follows [52]: CNNs establish connections in a localized manner, where each neuron interacts only with a small region of the input, rather than the entire preceding layer. This approach minimizes the number of parameters, improving efficiency and faster training.

These advantages make CNNs particularly effective for image recognition, object detection, and pattern analysis tasks.

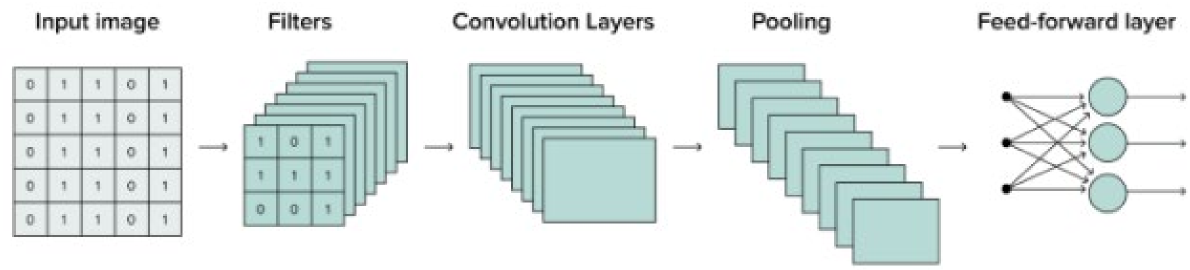


Figure 4. Visual Representation of a Common CNN Architecture

The convolution process applies a matrix, known as a feature detector or kernel, to the input data. This matrix contains predefined values that interact with the input to extract meaningful features. Adjusting the values in the feature detector generates different transformed versions of the input image, enabling the extraction of diverse features. We use standard convolutional layers with a stride of 1 and 'same' padding to preserve spatial dimensions [52] as demonstrated in Figure 5.

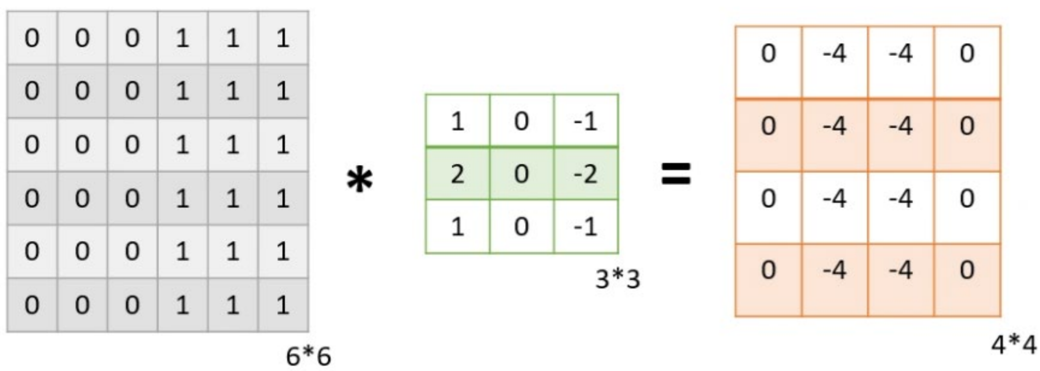


Figure 5. Padding in CNN

Transfer learning is a powerful concept in DL that leverages knowledge gained from one task (the "source task") to improve learning performance or reduce training time for a different but related task (the "target task"). In traditional ML, models are trained from scratch for each new task, which can be resource-intensive, especially for tasks requiring large datasets and complex architectures [53]. Transfer learning, however, allows for transferring learned patterns, representations, and weights from one model to another, significantly improving efficiency. The working principle of transfer learning is summarized step by step as follows:

Pre-trained Model as the Source: Transfer learning begins with a pre-trained model trained on a large, generic dataset such as ImageNet, which contains millions of labeled images across various classes [54].

Feature Extraction: The lower layers of neural networks, which generally capture basic, generic features like edges and textures, are often retained from the pre-trained model. These layers hold fundamental knowledge that is useful across tasks.

Fine-Tuning: The higher layers responsible for learning task-specific features are either retrained or replaced entirely to adapt the model to the target task. Fine-tuning can involve [55].

Freezing Layers: In scenarios where computational efficiency is prioritized, some layers are "frozen," meaning their weights are not updated during training, reducing computation time and data requirements.

Full Model Tuning: Alternatively, all layers may be adjusted on the new dataset, but with a lower learning rate to preserve the foundational knowledge while refining specifics.

Domain Adaptation: Some approaches to transfer learning also incorporate domain adaptation, a process that adjusts model features to account for differences between the source and target domains (e.g., a model trained on daytime images being adapted to nighttime images).

3.3 Experimented Transfer Learning Methods

All experiments were conducted on a workstation running Microsoft Windows 11 Pro with an Intel® Core™ i5-11400H CPU (2.70 GHz, 6 cores), 16 GB RAM, and an NVIDIA® GeForce RTX 3060 GPU (6 GB VRAM). The models were implemented in Python 3.8, using TensorFlow 2.11, Keras 2.11 for transfer learning architectures, and Scikit-learn 1.1.0 for classical classifiers.

Input images (256 × 256 × 3 RGB) were normalized and fed directly into the convolutional base of each pre-trained network without manual feature extraction. We fine-tuned five architectures, which are ResNet50, EfficientNetB0, InceptionV3, DenseNet121, and Xception, by replacing their top classification layers with a global average pooling layer followed by a dense layer of four neurons (softmax activation) corresponding to our classes (healthy, black rot, rust, scab).

Each transfer learning model was trained for 30 epochs with a batch size 32, using the Adam optimizer (learning rate = 1e-4) and categorical cross-entropy loss. Training requires approximately 3 hours per model. For comparison, the five classical ML models (SVM, Decision Tree, Random Forest, Logistic Regression, Naïve Bayes) were trained on flattened image vectors (196,608 features) using default hyperparameters and completed in under 20 minutes each. This clear distinction between input representations highlights the efficiency gains of transfer learning in processing raw image data without manual feature engineering.

3.3.1 ResNET

As neural networks deepen, challenges like the vanishing gradient problem are more likely to arise, where gradients diminish during backpropagation, making it difficult to update weights effectively. To address this, residual learning in CNNs revolutionized deep network design, as ResNet exemplified [56].

Residual learning uses residual blocks to help with the vanishing gradient problem. A residual block has a shortcut link over one or more layers. This lets the network learn residual functions instead of direct mappings. This method makes the optimization process easier and lets very deep networks be trained more quickly.

Figure 6 shows how the architecture of a three-layer residual block shows how these shortcuts help gradient flow and keep performance high as the network depth increases.

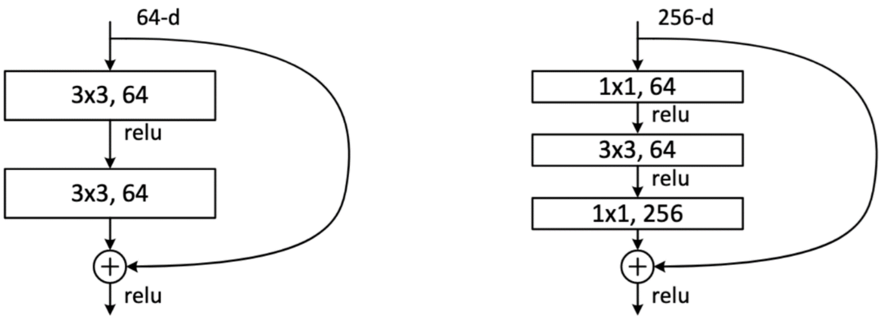


Figure 6. Residual Block Samples of ResNET with RELU as Activation Function

3.3.2 EfficientNET

The Mobile Inverse Bottleneck (MB) convolution block is a key component of EfficientNet, known for its fast speed. This structure facilitates the accurate interpretation and processing of high-level and spatial features and works well for downscaling during segmentation tasks. EfficientNet uses a novel connection method to improve feature representation without sacrificing accuracy [57].

The new layer in EfficientNet features a thin bottleneck layer that can assist in tasks such as identification, image classification, and segmentation. In particular, the inclusion of residual connections in segmentation significantly enhances functionality. This simplifies the collection and presentation of additional features. Batch normalization and activation functions are applied after each convolution operation. This reduces issues such as vanishing gradients and ensures stable training.

The inclusion of residual blocks in EfficientNet has shown improvements in segmentation metrics across a variety of datasets. This architecture makes it easier for the network to identify complex features, improving the accuracy of pixel-by-pixel segmentation tasks. These changes demonstrate the adaptability and reliability of EfficientNet in achieving superior results in computer vision applications.

3.3.3 Inception

The InceptionV3 model is well-known for fast data processing because it uses factorization to split inputs into large and small groups [58]. This method uses the least number of resources possible while still being accurate.

The InceptionV3 architecture has layers that work on input data simultaneously. These filters come in different sizes. After that, the results of these filters are combined to make a single feature map. This design lets the model look at the same data from different angles and find connections between channels and locations [59].

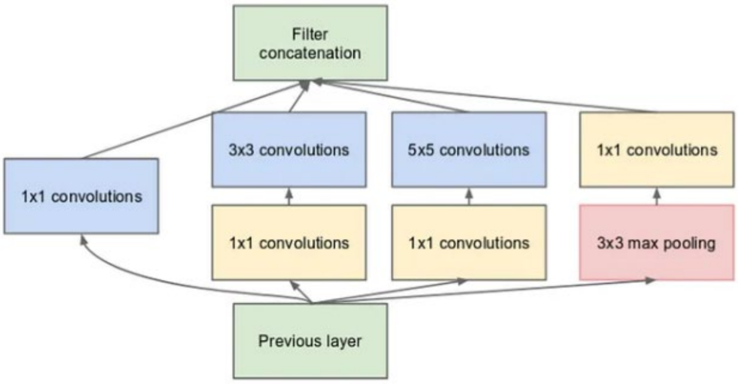


Figure 7. Example of Inception v3 Model Block

3.3.4 DenseNet

DenseNet is a CNN architecture designed to maintain high performance and improve feature reuse. It also reduces the number of trainable parameters. This occurs through a dense connection, where each layer is connected to all other layers in a feedforward manner [60]. Subsequent layers can directly access features learned in previous layers, reducing redundancy and improving gradient flow during training.

The model uses a lightweight architecture consisting of a transition layer and two parallel dense blocks, as described. This streamlined design reduces computational complexity and makes learning efficient.

Originally designed for image classification, DenseNet was developed to process one-dimensional voltage-current time series data. This change increases the n×n convolution kernel size to 1×n, allowing the model to process sequences rather than spatial data. This configuration allows the network to effectively capture temporal features in error detection tasks or similar applications. This configuration optimally balances robust feature extraction with a lightweight architecture, meeting domain-specific requirements while being ideal for resource-constrained environments.

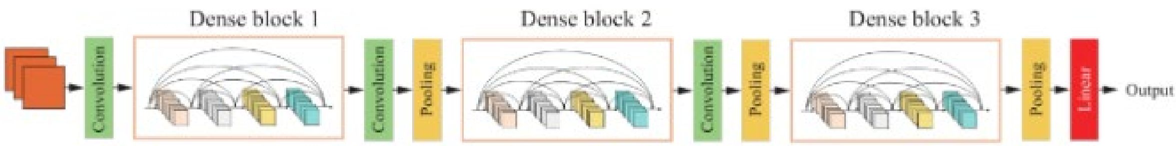


Figure 8. DenseNET Overall Structure

3.3.5 Xception

This model aims to determine the level of driver distraction by analyzing each frame. It converts images using textual labels that conform to predefined categories for classification and discrimination [61]. The Xception model helps identify body components and regions of interest (ROIs) associated with distracting activities. Precise predictions can be made thanks to new technology. The Xception model can extract convolutional features from the input image. These features capture complex patterns and relevant details in the image. The model groups each frame into specific distraction labels. It identifies distracting objects or elements in the image, such as mobile devices or movements. Localizing relevant body parts or actions within the frame constitutes a distraction. The model provides the final predictions and distraction analysis by combining the classification and detection results.

4. Experimental Studies

This section summarizes the prediction accuracies, recall rates, and f-scores obtained by the implemented methods. Traditional and transfer learning models used a dataset containing apple leaf images. As mentioned earlier, the data class labels are reduced to two classes: healthy and diseased leaf images. This preprocessing step converts all unhealthy class labels, including rust, black rot, and scab, into the diseased class. Consequently, the dataset becomes much more balanced since the number of instances of each class is nearly equal. In the first part of the experiments, the dataset was randomly divided into two parts: training and testing. 80% of the entire dataset is the training set, while the remaining 20% is the testing set.

At this stage, 5 different conventional ML methods have been applied to classify apple plant leaves as healthy or diseased. These methods are support vector machine (SVM), decision trees (DT), random forest (RT), logistic regression (LR), and Naïve Bayes (NB) classifiers. SVM obtained 0.71 F1-Score and 0.64 accuracy values. Table 3 reveals the classification performance results with accuracy, precision, recall, and F-score values.

Table 3. Classification Performance Results of 5 Conventional ML Algorithms Applied to Apple Plant Leaf Dataset Divided into Two Parts: 80% Training and 20% Testing Sets

Model	Accuracy	F1 Score	Precision	Recall
SVM	0.6445	**0.7198**	0.6188	**0.8601**
DT	0.5024	0.5388	0.5303	0.5476
RF	0.6082	0.6342	0.6287	0.6399
LR	**0.6524**	0.6901	**0.6551**	0.7292
NB	0.6240	0.6657	0.6303	0.7054

As can be seen in Table 3, LR has the highest accuracy and precision values, while SVM obtains the best recall and F1-Score values. DT acquires far worse performance values than other methods. All of these five models have been implemented in the second part of the experiment, but the dataset is split into two parts using a 10-fold cross-validation technique. The testing part is selected randomly as 10% of the dataset. In each ten iterations, 10% of the data becomes the testing set, and the remaining instances become the training set. Classification is performed in each iteration, and finally, the average results are recorded as classification performance values of the applied classifier. As a result, Table 4 shows the outcomes of the same five conventional classifiers for the 10-fold cross-validation process. It can be observed that there is no significant difference in the results compared to Table 3.

Table 4. Classification Performance Results of 5 Conventional ML Algorithms Applied to Apple Plant Leaf Dataset Divided into Two Parts using 10-Fold Cross-Validation

Model	Accuracy	F1 Score	Precision	Recall
SVM	**0.6317**	**0.7072**	0.6014	**0.8584**
DT	0.5078	0.5297	0.5248	0.5350
RF	0.5872	0.6159	0.5951	0.6401
LR	**0.6317**	0.6648	**0.6292**	0.7047
NB	0.6008	0.6412	0.6004	0.6882

In the third part of the experiment, a 10-fold cross-validation was performed first. Then, five different transfer learning models mentioned in Section 3 have been implemented. Although several versions of these techniques exist, only one version has been selected for those models. These methods are ResNet50, EfficientNetB0, InceptionV3, DenseNet121, and Xception. In Table 5, classification performance values for these models are shown. It implies a great improvement for each model, so it can be concluded that the transfer learning mechanism is much more suitable for plant leaf disease detection, especially in image processing.

For the deep learning models, a stratified 10-fold cross-validation strategy was applied to the entire dataset (3,164 images). In each fold, approximately 90% of the data (~2,847 images) was used for training and 10% (~317 images) for validation, with class distributions preserved. The dataset was split into training and testing subsets for the machine learning models using a stratified 80–20 ratio, resulting in 2,531 images for training and 633 for testing. This consistent stratification ensured that each subset maintained the original class balance. The use of cross-validation in deep learning and fixed train–test splitting in machine learning allowed for a robust comparison between the two model groups while minimizing the impact of random sampling bias.

Table 5. Classification Performance Results of 5 Transfer Learning Algorithms Applied to Apple Plant Leaf Dataset Divided into Two Parts Using 10-Fold Cross-Validation

Model	Accuracy	F1 Score	Precision	Recall
ResNet50	0.9953	0.9954	0.9939	0.9970
EfficientNetB0	**0.9994**	**0.9994**	**0.9994**	**0.9994**
InceptionV3	0.9264	0.9438	0.9107	0.9956
DenseNet121	0.9302	0.9124	0.9980	0.8686
Xception	0.9991	0.9991	0.9988	**0.9994**

Table 5 indicates that although all transfer learning approaches present highly acceptable improvements compared to traditional ML classifiers, EfficientNet can be considered the most reliable model for apple plant disease detection tasks using leaf images. It acquires more than 0.99 value for all of the performance metrics. Thus, it outperforms the other four transfer learning methods in each metric.

EfficientNetB0 achieves higher accuracy with fewer parameters due to its compound scaling strategy, which uniformly balances network depth, width, and input resolution. Its use of mobile inverted bottleneck convolution (MBConv) blocks with squeeze-and-excitation modules enables more efficient feature extraction per parameter, outperforming larger architectures such as ResNet50.

Additionally, Figure 9 gives information about the number of correctly classified and misclassified instances by EfficientNet. It presents a confusion matrix for this model. It shows a visualization of predicted values for true positives, false negatives, false positives, and true negatives across each class label. The confusion matrix reveals high diagonal values (306 and 326) for the respective classes, with minimal non-diagonal elements, demonstrating the model's strong performance and reliability in detecting plant leaf diseases. Out of 632 predictions, EfficientNet produced only three incorrect classifications, underscoring its accuracy and effectiveness.

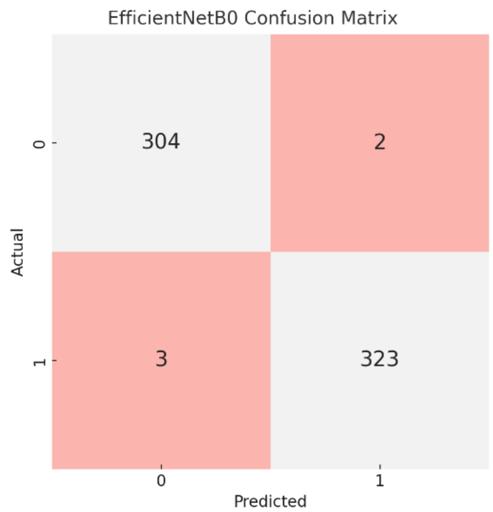


Figure 9. Confusion matrix of EfficientNet

5. Conclusion and Future Work

Developing detection systems is essential for achieving accurate, consistent, and reliable predictions in agriculture. Plant disease detection can be accomplished using traditional image processing techniques or by employing DL algorithms, which can be incorporated at various stages of a plant's lifecycle. Automation technologies in agriculture can significantly simplify and streamline the analysis and identification of plant diseases. These advancements promote a healthier and more sustainable environment for plants and help mitigate significant risks, such as the potential extinction of various plant species.

We evaluated five classical ML methods (Support Vector Machine, Decision Tree, Random Forest, Logistic Regression, Naïve Bayes) alongside five deep transfer learning architectures (ResNet50, EfficientNetB0, InceptionV3, DenseNet121, and Xception). Among the transfer learning models, EfficientNetB0 achieved the highest accuracy (99.93%), closely followed by Xception (99.91%) and ResNet50 (99.53%). DenseNet121 and InceptionV3 also demonstrated strong results, with accuracies of 93.02% and 92.64%, respectively. These results confirm that transfer learning methods outperformed traditional ML approaches, which achieved maximum accuracies of approximately 65% (Logistic Regression) to 64% (SVM). This comparative analysis highlights EfficientNetB0 as the most reliable option for practical deployment, while also showcasing the competitive performance of other state-of-the-art architectures.

The EfficientNetB0 model achieved the highest accuracy rate of 99.93% compared to other models. This high performance aligns with its efficient architecture, rapid computational capability, and exceptional success in transfer learning. ResNet50 and Xception models also delivered highly successful results, achieving remarkably high accuracy and recall rates.

Although the dataset contains 3,164 apple leaf images, which is relatively small for training DL models, potential biases must also be considered. The images originate from a single publicly available dataset, which may limit diversity in lighting conditions, camera angles, and orchard environments. This homogeneity can restrict the generalizability of the trained model to real-world scenarios. To mitigate these limitations, several augmentation strategies (e.g., random rotation, brightness adjustment, horizontal/vertical flipping, and scaling) were employed during training to expand data diversity artificially. However, future studies should incorporate additional real-field images collected under various weather conditions and geographic locations to minimize sampling bias further and enhance robustness [62].

The apple dataset utilized in this study includes four distinct categories of leaves. Concentrating on a smaller number of classes enables a more targeted, practical, and comprehensive analysis of these specific categories, facilitating the identification of their characteristics, underlying causes, and effective mitigation strategies. In certain agricultural contexts, prioritizing the detection of a few prevalent diseases may be more relevant and immediately beneficial for farmers than a broader classification approach that encompasses a wide range of diseases. This proposed methodology presents numerous opportunities for future research, from enhancing model accuracy to real-world applications, ultimately supporting more efficient and profitable farming practices.

Early and accurate detection of leaf diseases enables targeted pesticide application to affected trees, thereby significantly reducing overall chemical use [63]. Moreover, by feeding time-stamped detection data into crop health models, farmers can forecast disease outbreaks and proactively adjust irrigation or nutrient management. However, over-reliance on AI predictions may lead to misdiagnoses under novel environmental conditions; we recommend integrating human expert review and periodic model retraining with new field data.

Several practical deployment concerns must be acknowledged, even though this study proposes using mobile applications to assist farmers. To begin with, DL models often require high computational resources, which may limit their scalability on low-end mobile devices with limited memory and processing capabilities. To overcome this, model compression or lightweight architecture (e.g., MobileNet) can be explored in future implementations. Next, limited internet connectivity and smartphone penetration may hinder access to such tools in rural or low-resource regions. Moreover, varying levels of digital literacy among farmers also present difficulties, necessitating intuitive user interfaces, multilingual support, and possibly offline functionality. These factors should be considered when adapting AI-based solutions for real-world agricultural settings, ensuring that technological innovations are usable and equitable.

Socioeconomic factors must also be addressed for equitable distribution. This should go beyond computational limitations. The use of AI-based disease detection systems can be hindered in many rural areas by limited internet connectivity, low smartphone usage, and farmers' digital literacy. Future implementations should consider offline inference modes, compatibility with low-cost hardware, and an intuitive user interface with multilingual support to address these issues. Furthermore, pilot studies involving end users can ensure the developed system is compatible with real-world agricultural workflows and prevent it from inadvertently exacerbating technology disparities between high- and low-resource agricultural communities.

We may encounter several challenges when applying our EfficientNetB0-based model to resource-constrained platforms such as agricultural drones or mobile devices. First, mobile CPU/GPUs have limited processing capacity and memory, which can increase inference latency. Model optimization techniques such as pruning, quantization, and information distillation should be used to address this issue. Second, drones' battery life and energy consumption limit continuous operation. Lightweight runtimes such as TensorFlow Lite or ONNX Runtime Mobile are recommended for efficient execution. Third, in orchards with unreliable or low-bandwidth network connections, all dependencies may need to be consolidated locally. Finally, motion blur and changing camera angles can affect real-time image capture on moving platforms. Therefore, built-in preprocessing such as frame stabilization and blur detection should be incorporated into the software pipeline.

Although the model achieved 99.93% accuracy under controlled conditions, performance may degrade when applied to real-world orchard environments where lighting, background clutter, and leaf orientation vary significantly. Future studies should incorporate domain adaptation strategies, including data collection from multiple orchards across different seasons and using augmented samples (e.g., brightness jitter, perspective transforms, random rotations) to simulate field variability. Additionally, periodic retraining with newly acquired field data can enhance robustness and ensure sustained accuracy during deployment.

References

[1] H. S. Harshitha, J. Nagaraja, and D. Pruthiraja, “Plant disease detection using image processing,” Proc. - 2024 Second International Conference on Advances in Information Technology (ICAIT), pp. 1–6, 2024.

[2] M. P. Taranukho, Y. M. Kovalyshyna, and Y. V. Zaika, “Effect of viral infection on the ultrastructural organization of blackcurrant leaf tissue cells,” Mikrobiolohichnyi Zhurnal, vol. 84, no. 5, pp. 38, 2022.

[3] R. Yakkundimath, G. Saunshi, and S. Palaiah, “Automatic methods for classification of visual-based viral and bacterial disease symptoms in plants,” Springer, 2021.

[4] A. A. Yatoo and A. Sharma, “An indigenous dataset for the detection and classification of apple leaf diseases,” Data in Brief, vol. 53, p. 110165, 2024.

[5] S. Kumar, R. Kumar, M. Gupta, and A. J. Obaid, “EEDL-based detection and classification of apple foliar leaf disease,” Proc. - 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–6, 2024.

[6] S. Huo, N. Duan, and Z. Xu, “An improved multi-scale YOLOv8 for apple leaf dense lesion detection and recognition,” IET Image Processing, 2024.

[7] J. C. Semenza, J. Rocklöv, and K. Ebi, “Climate Change and Cascading Risks from Infectious Disease,” Infectious Diseases and Therapy, vol. 11, 2021.

[8] J. K. Chavda and B. D. Vaniya, “Plant disease identification and classification using machine learning models: A survey,” Proceedings of IEEE Smart Technologies, pp. 281–289, 2023.

[9] S. Sharma, A. Kapoor, and A. Jain, “Analysis of plant diseases using image processing techniques,” International Journal of Computer Vision, vol. 128, no. 3, pp. 343–362, 2023.

[10] M. Chowdhury, M. O. Rahman, and S. Alam, Proprietor: “A farmer assistance smartphone application with crop planner, crop disease help, agri-expert search, and crop suggestion features,” 15th International Conference on Computing Communication and Networking Technologies, ICCCNT 2024, pp. 1–6, 2024.

[11] P. Kambale, D. R. B. M., D. Patil, and R. G. N., “Mobile technology for farmers: An overview of agricultural apps,” Asian Journal of Agricultural Extension, Economics & Sociology, vol. 42, no. 9, pp. 75–81, 2024.

[12] I. Kunduracioglu, “CNN models approaches for robust classification of apple diseases,” CNN Models Approaches for Robust Classification of Apple Disease, vol. 1, pp. 235–251, 2024.

[13] P. Matov, S. Filipova-Petrakieva, M. Lazarova, and I. Taralova, “Apple trees leaves pathologies detection using deep learning convolutional neural network,” 2024 IEEE International Conference on Omni-Layer Intelligent Systems, COINS 2024, pp. 1–4, 2024.

[14] N. Dembla and R. Yadav, “Advancing foliar disease detection in apple orchards: Evaluating the efficacy of the EfficientNetB3 model in machine learning-based classification,” 2024 International Conference on Advances in Computing Research on Science, Engineering and Technology, ACROSET 2024, pp. 1–4, 2024.

[15] Z. Qiu, Y. Xu, C. Chen, W. Zhou, and G. Yu, “Enhanced disease detection for apple leaves with rotating feature extraction,” Agronomy, vol. 14, no. 11, pp. 2602–2602, 2024.

[16] R. W. Mwangi, M. Mustafa, K. Charles, I. W. Wagara, and N. Kappel, “Selected emerging and reemerging plant pathogens affecting the food basket: A threat to food security,” Journal of Agriculture and Food Research, vol. 14, 2023.

[17] A. Awari, V. Bhokare, H. Daundkar, S. Dhas, and Prof. A. Take, “Plant disease detection and classification,” Indian Scientific Journal of Research in Engineering and Management, vol. 08, no. 10, pp. 1–7, 2024.

[18] R. K. Ramya and C. Meenakshi, “A review on ML and DL techniques in detecting plant diseases,” International Journal

of Advanced Research in Science, Communication and Technology, pp. 34–40, 2024.

[19] D. Navaneetha, T. Maringanti, S. Repaka, V. Basalingwar, Y. Vasavi, and N. V. Manasa, “Analysis of automatic health monitoring in agricultural crops using artificial intelligence methods,” 2nd IEEE International Conference on Data Science and Network Security, ICDSNS 2024, pp. 1–4, 2024.

[20] S. Arulmurugan, V. Bharathkumar, S. Gokulachandru, and M. Mohamad Yusuf, “Plant guard: AI-enhanced plant diseases detection for sustainable agriculture,” 7th International Conference on Inventive Computation Technologies, ICICT 2024, pp. 726–730, 2024.

[21] A. G. Khan, R. Javed, A. M. Malik, A. F. Mirza, S. Aslam, A. Shabbir, and H. Tauseef, “Early detection and classification of apple leaf diseases using deep learning technique,” Deleted Journal, vol. 4, no. 3, 2024.

[22] P. S. Soltis, L. Teixeira-Costa, P. Bonnet, and R. G. Nelson, “Advances in plant imaging across scales,” Applications in Plant Sciences, vol. 11, no. 5, 2023.

[23] A. Sulaiman, V. Anand, S. Gupta, H. Alshahrani, M. S. Al Reshan, A. Rajab, A. Shaikh, and A. T. Azar, “Sustainable apple disease management using an intelligent fine-tuned transfer learning-based model,” Sustainability, vol. 15, no. 17, 2023.

[24] K. Sujatha, K. Gayatri, M. S. Yadav, N. C. Sekhara Rao, and B. S. Rao, “Customized deep CNN for foliar disease prediction based on features extracted from apple tree leaves images,” International Interdisciplinary Humanitarian Conference for Sustainability, IIHC 2022 - Proceedings, pp. 193–197, 2022.

[25] P. Matov, S. Filipova-Petrakieva, M. Lazarova, and I. Taralova, “Apple trees leaves pathologies detection using deep learning convolutional neural network,” 2024 IEEE International Conference on Omni-Layer Intelligent Systems, COINS 2024, pp. 1–4, 2024.

[26] S. Alqethami, B. Almtanni, W. Alzhrani, and M. Alghamdi, “Disease detection in apple leaves using image processing techniques,” Engineering, Technology & Applied Science Research, vol. 12, no. 2, pp. 8335–8341, 2022.

[27] T. Biswas, S. K. Pati, and S. Sarangi, “Plant disease detection,” Prospects of Science, Technology and Applications, pp. 250–256, 2024.

[28] S. Gupta, S. Jain, and K. C. Bandhu, “Revolutionizing plant disease detection: A comprehensive review of machine learning techniques and applications,” 2024 International Conference on Advances in Computing Research on Science Engineering and Technology, ACROSET 2024, pp. 1–7, 2024.

[29] P. Patwal, R. Chauhan, C. Bhatt, and S. Devliyal, “Automated tomato disease detection and classification using image processing and machine learning for precision agriculture,” Challenges in Information, Communication and Computing Technology, pp. 481–486, 2024.

[30] V. V. N. Vinaykumar, A. Babu, P. Bhargavi, and S. Aluvala, “An Inception Architecture with Transfer Learning for Apple Plant Leaf Disease Detection,” in Proceedings of the 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE), SR University, Nov. 2023. DOI: 10.1109/AIKIIE60097.2023.10390058.

[31] Khan, A., & Samantaray, S. D. Comparative Analysis of Pre-Trained CNN Architectures for Apple Foliar Disease Classification. Indian Scientific Journal Of Research In Engineering And Management, 2023. https://doi.org/10.55041/ijsrem25943

[32] Rawat, P., & Singh, S. K. Apple Leaf Disease Detection Using Transfer Learning. 1–6, 2024. https://doi.org/10.1109/icicacs60521.2024.10498746

[33] Babalola, F. O., Kpai, N. I., & Toygar, Ö. Elma Yaprağı Hastalıklarının AlexNet Kullanılarak Derin Öğrenme Tabanlı Sınıflandırılması. Computer Science, 2023. https://doi.org/10.53070/bbd.1349566

[34] Ashmafee, Md. H., Ahmed, T., Ahmed, S., Hasan, Md. B., Jahan, Mst. N., & Rahman, A. An Efficient Transfer Learning-based Approach for Apple Leaf Disease Classification. European Conference on Cognitive Ergonomics, 2023, 1–6. https://doi.org/10.1109/ECCE57851.2023.10101542

[35] Saharan, M. S., & Singh, G.. Leaf Disease Detection Using Transfer Learning (pp. 44–58). Springer Science+Business Media, 2023. https://doi.org/10.1007/978-3-031-47997-7_4

[36] M. H. Rabbi, M. J. Hossain, A. Masud, and A. Rahman, “An image recognition approach for crop disease detection in agro-field from infected plant area,” Global Mainstream Journal of Innovation, Engineering & Emerging Technology, pp. 11–18, 2024.

[37] I. Ahmed and P. K. Yadav, “A systematic analysis of machine learning and deep learning based approaches for identifying and diagnosing plant diseases,” Sustainable Operations and Computers, vol. 4, pp. 96–104, 2023.

[38] T. T. Moharekar and U. Pol, "Detection and classification of plant leaf diseases using convolution neural networks and Streamlit," www.irjmets.com, 2022.

[39] V. Swetha and R. Jayaram, "A novel method for plant leaf malady recognition using machine learning classifiers," Proceedings of the 3rd International Conference on Electronics and Communication and Aerospace Technology, ICECA 2019, pp. 1360–1365, 2019.

[40] Y. Wang, Y. Wei, X. Yu, J. Wang, Y. Zhang, L. Zhang, Y. Wan, and Z. Chen, "A segmentation network for generalized lesion extraction with semantic fusion of transformer with value vector enhancement," Expert Systems with Applications, vol. 266, p. 126098, 2025.

[41] N. Shinde and A. Ambhaikar, "Comprehensive survey on datasets, models, and future directions in plant disease prediction," International Journal of Image and Graphics, 2024.

[42] Y. He, Q. Gao, and Z. Ma, "A crop leaf disease image recognition method based on bilinear residual networks," Mathematical Problems in Engineering, vol. 2022, p. 2948506, 2022.

[43] Y. A. Nanehkaran, D. Zhang, J. Chen, Y. Tian, and N. Al-Nabhan, "Recognition of plant leaf diseases based on computer vision," Journal of Ambient Intelligence and Humanized Computing, 2020.

[44] M. Bhagat and D. Kumar, "Efficient feature selection using BoWs and SURF method for leaf disease identification," Multimedia Tools and Applications, vol. 82, no. 18, pp. 28187–28211, 2023.

[45] A. Saygılı, "The efficiency of transfer learning and data augmentation in lemon leaf image classification," European Journal of Engineering and Applied Sciences, vol. 6, no. 1, pp. 32–40, 2023.

[46] B. Dikici, M. F. Bekçioğullari, H. Açikgöz, And D. Korkmaz, "Performance investigation of pre-trained convolutional neural networks in olive leaf disease classification," Konya Journal of Engineering Sciences, vol. 10, no. 3, pp. 535–547, 2022.

[47] Y. Kurmi and S. Gangwar, "A leaf image localization based algorithm for different crops disease classification," IPAgr, vol. 9, no. 3, pp. 456–474, 2022.

[48] M. Türkoğlu, K. Hanbay, I. Sivrikaya, and D. Hanbay, "Derin evrişimsel sinir ağı kullanılarak kayısı hastalıklarının sınıflandırılması," Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, vol. 9, no. 1, pp. 334–345, 2020.

[49] Cybenko, G. Neural networks in computational science and engineering. 3(1), 36–42, 1996. https://doi.org/10.1109/99.486759

[50] Rose, A. How do Artificial Neural Networks Work. Journal of Advances in Science and Technology, 20(1), 172–177, 2024. https://doi.org/10.29070/ttrkmm98

[51] Haug, C. Convolutional Neural Networks (pp. 139–176). Apress eBooks, 2022. https://doi.org/10.1007/978-1-4842-8835-1_3

[52] S. A. Wagle and R. Harikrishnan, "Comparison of plant leaf classification using modified alexnet and support vector machine," Traitement Du Signal, vol. 38, no. 1, pp. 79–87, 2021.

[53] Z. Chen, H. Qin, W. Xiang, and Y. Liu, "Reshaping the feature distribution to aid in building extraction," Computer Vision and Pattern Recognition, p. 27, 2024.

[54] L. Torrey and J. Shavlik, "Transfer learning," in 978-1-60566-766-9.CH011, pp. 242–264, 2010.

[55] M. Zoric, M. Stula, I. Markic, and M. Braovic, "Transfer learning in building neural network model case study," 2024 32nd International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2024, pp. 1–6, 2024.

[56] V. Gunturu, N. Maiti, B. Toure, P. Kunekar, S. B. Banu, and D. Sahaya Lenin, "Transfer learning in biomedical image classification," in Proceedings - 3rd International Conference on Advances in Computing, Communication and Applied Informatics, ACCAI 2024, 2024.

[57] N. Pateriya, P. Jain, K. P. Niveditha, V. Tiwari, and S. Vishwakarma, "Deep residual networks for image recognition," International Journal of Innovative Research in Computer and Communication Engineering, vol. 11, no. 09, pp. 10742–10747, 2023.

[58] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," PMLR, pp. 6105–6114, 2019.

[59] X. Xia, C. Xu, and B. Nan, "Inception-v3 for flower classification," in 2017 2nd International Conference on Image, Vision and Computing, ICIVC 2017, pp. 783–787, 2017.

[60] G. Huang, S. Liu, L. van der Maaten, and K. Q. Weinberger, "CondenseNet: An efficient DenseNet using learned group convolutions," pp. 2752–2761, 2018.

[61] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," pp. 1251–1258, 2017.

[62] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 1, pp. 1–48, 2019.

[63] A. G. Khan, R. Javed, A. M. Malik, A. Faisal, S. Aslam, A. Shabbir, and H. Tauseef, "Early Detection and Classification of Apple Leaf Diseases Using Deep Learning Technique," The Asian Bulletin of Big Data Management, vol. 4, no. 3, Sep. 2024, doi:10.62019/abbdm.v4i3.217.

Article Information Form

Authors Contributions

Cemal Yüksel: review and editing (equal). Alican Doğan: Conceptualization (lead); writing – original draft (lead); formal analysis (lead); writing – review and editing (equal)- Software (lead).

Acknowledgments

This work has not been financially funded by any sponsor or institution. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

Conflict of Interest Notice

The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical Approval and Informed Consent

This study utilized publicly available datasets containing apple leaf images and did not involve the collection of new data from human subjects or farmers. Therefore, no ethical approval or informed consent was required. All analyses adhered to recognized scientific and ethical standards, and the data usage complied with the original dataset's licensing terms.

Availability of data and material

Not applicable / or link

Artificial Intelligence Statement

The authors confirm that no artificial intelligence tools were employed in the preparation or authorship of this article.

Plagiarism Statement

This article has been scanned by iThenticate™.