International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Revue d'Intelligence Artificielle
Vol. 38, No. 2, April, 2024, pp. 717-727
Journal homepage: http://iieta.org/journals/ria

Deep Learning Approaches for Potato Leaf Disease Detection: Evaluating the Efficacy of Convolutional Neural Network Architectures

Erlin[1*], Indra Fuadi[2], Ramalia Noratama Putri[3], Dewi Nasien[1], Gusrianty[1], Dwi Oktarina[3]

[1] Department of Informatics Engineering, Institut Bisnis dan Teknologi Pelita Indonesia, Pekanbaru 28127, Indonesia
[2] Food Crop and Horticultural Protection Agency of Riau Province, Pekanbaru 28294, Indonesia
[3] Department of Information Systems, Institut Bisnis dan Teknologi Pelita Indonesia, Pekanbaru 28127, Indonesia

Corresponding Author Email: erlin@lecturer.pelitaindonesia.ac.id

https://doi.org/10.18280/ria.380236

Received: 11 December 2023
Revised: 1 January 2024
Accepted: 11 February 2024
Available online: 24 April 2024

Keywords:
agriculture, CNN architectures, computer vision, deep learning, detection, image classification, potato leaf disease, pre-trained model

ABSTRACT

In agriculture, timely and accurate detection of plant diseases is essential to obtain healthy crop yields and ensure food security. However, detecting diseases in potato leaves is challenging because of the complex symptoms and variability in leaf appearances. This requires the development of an effective and efficient method that can overcome these challenges and improve disease detection accuracy. Utilizing the power of computer vision and deep learning, this paper presents a comprehensive study on potato leaf disease detection using a multi-architecture Convolutional Neural Networks (CNNs) approach. We evaluate five different CNN architectures: VGG16, VGG19, MobileNetV2, ResNet50, and AlexNet, to assess their classification capabilities. The research encompassed the dataset collection, data augmentation, model selection, hyperparameter tuning, and evaluation, leading to a rigorous analysis of detection accuracy, model convergence, and training efficiency. Our findings revealed that ResNet50 was the standout performer, achieving a remarkable 97% testing accuracy and 98% specificity. Conversely, the VGG19 architecture was the least effective. A consistent challenge across all models was accurately classifying categories of healthy leaves, indicating a potential area for model refinement. This study not only highlights the efficacy of deep learning in plant health diagnosis but also highlights the importance of specificity as an important metric in such tasks. The results of our study provide a promising avenue for real-time diagnosis of potato leaf diseases in the field, paving the way for healthier crops and increased agricultural productivity.

1. INTRODUCTION

Potato (Solanum tuberosum) is a widely cultivated crop that has significant economic value globally. Its ability to adapt to diverse climates and its high nutritional value make it an important food source for millions of people [1]. Potatoes are rich in carbohydrates, vitamins, and minerals and play an important role in a balanced diet [2]. In addition, potatoes are also a staple food in some regions, especially in developing countries where they contribute to food security and livelihoods. However, the threat posed by potato foliar diseases is a major challenge facing farmers and researchers. Potatoes are susceptible to various diseases caused by pathogens such as fungi, bacteria, and viruses [3]. These diseases manifest through symptoms such as leaf discoloration, wilting, necrosis, and deformation, which ultimately affect the overall health and productivity of the crop. The impact of these diseases on crop yields can be severe, including reduced yields, reduced quality, increased production costs, and food safety concerns.

Detecting and treating potato leaf diseases at an early stage is crucial to maintaining crop vitality. Traditional methods of identifying diseases often involve manual inspection, which can be time-consuming, subjective, and prone to errors.

The integration of advanced technologies and artificial intelligence has introduced innovative solutions. One such breakthrough technology is Convolutional Neural Networks (CNNs), which is revolutionizing the field of plant foliar disease detection. CNNs offer a transformative approach by harnessing the power of deep learning to improve the accuracy, efficiency, and scalability of disease detection.

CNNs are a class of deep learning models specifically designed for processing and analyzing visual data. Inspired by the human visual system, they excel at recognizing complex patterns and features in images. This unique ability makes CNNs well-suited for the task of detecting plant leaf diseases [4]. By training a diverse dataset of images containing healthy and diseased leaves, CNNs can learn to identify subtle visual cues that indicate the presence of disease [5, 6].

The successful application of CNNs in detecting potato leaf diseases based on images is closely linked to several variables, one of the most important being the choice of architecture. Some popular architectures are ResNet [7], VGG [8], Inception [9], AlexNet [10], Squeezenet [11], EfficientNet [12], DenseNet [13], and MobileNet [14]. ResNet, short for Residual Network, is a CNN architecture designed to address the missing gradient problem, which is a common problem in deep learning associated with increasing the number of layers. The ResNet architecture uses jump connections between layers to add the output of the previous layer to the output of the stacked layers, thus resulting in the ability to train a much deeper network than was previously possible [15]. The VGG (Visual Geometry Group) architecture, specifically VGG16, and VGG19, is a classic convolutional neural network architecture known for its simplicity and effectiveness in various computer vision tasks [16]. The Inception architecture, also known as GoogLeNet, was developed to address the challenge of increasing the depth and width of CNNs while managing computational complexity [17]. Squeezenet is a popular CNN architecture designed to create smaller neural networks with fewer parameters, allowing for easier fitting into computer memory and transmission over networks [18]. EfficientNet, introduced by a research team at Google, was described in the study [19] as a new method for scaling CNN architectures. Furthermore, DenseNet is a type of CNNs that exploits dense connections between layers through dense blocks, where all layers (with appropriate feature map sizes) are directly adjacent to each other [20]. Finally, MobileNet is designed to be computationally efficient and suitable for mobile devices and embedded devices thus making it a promising approach for automated farming [21].

Although many studies have explored the use of convolutional neural network architectures, there seems to be a limited comparative analysis of these architectures specifically for potato leaf disease classification. The optimal architecture for accurate and efficient potato leaf disease classification is still unclear. Moreover, the variations in disease type, severity, and image quality in potato leaf disease datasets present unique challenges that may require customized CNN architectures.

This research provides valuable insight into the most effective CNN architectures to determine which one is best suited for a given disease detection task. This could include comparing accuracy, computational efficiency, and other relevant metrics, taking into account the nuances of the dataset and the specific characteristics of potato diseases. This research not only contributes to the broader field of potato leaf disease classification but also provides practical guidance for agricultural practitioners seeking better disease detection and management strategies for potato production.

The paper is organized as follows: Section 2 presents a comprehensive review of the relevant literature, providing the historical context of previous research. Section 3 describes the materials and methods used to detect potato leaf diseases. Section 4 presents the results and discussion. Finally, in Section 5, we summarize the paper with concluding remarks and present future research.

2. RELATED WORK

The importance of automating plant disease diagnosis has been increasingly recognized in the agricultural domain. Early work in this area mainly used traditional machine learning techniques. Hylmi et al. [22] conducted a study on the detection of potato leaf disease using a multi-class support vector machine based on texture, color, and shape features. The study, which began with leaf spot extraction and used RGB histograms, GLCM, and contour calculations, achieved a 97.56% accuracy rate. Another study used K-Means clustering to enhance potato production and the model achieved a 97% accuracy rate [23]. Rusli et al. [24] proposed a potato leaf disease classification technique using image processing and artificial neural network (ANN) methods, aimed at early detection of plant diseases to reduce losses in agricultural production. Another study introduced an image processing and machine learning-based system to identify and classify potato leaf diseases, specifically Early Blight (EB) and Late Blight (LB), crucial for enhancing crop production in Bangladesh. By utilizing image segmentation on 450 images from the Plant Village database and employing seven classifier algorithms, the study found the Random Forest classifier to be the most effective, achieving a 97% accuracy rate in distinguishing diseased and healthy leaves [25].

These traditional methods often require more manual input and are not as efficient as machine learning-based approaches. Additionally, these methods often struggle with adapting to data variations and typically require more frequent adjustments and maintenance [26].

On the other hand, the emergence of Deep Learning and specifically Convolutional Neural Networks has revolutionized this field. CNNs have shown significant advantages in capturing hierarchical features in images, thus eliminating the need for manual feature extraction. Some studies have provided a comprehensive overview of the application of CNNs for plant disease classification, highlighting the advantages of deep learning architectures over traditional methods [27-31].

Over the years, several deep CNN architectures have been proposed and compared in the context of plant disease classification. VGG-16 is a convolutional neural network architecture that has been used in several studies to detect plant diseases with high accuracy [32]. The VGG-16 model was used to detect 19 different classes of plant diseases with an accuracy of about 95.2% [33]. Another study used transfer learning with VGG-16 to classify multi-plant leaf disease images with high accuracy [34]. In addition, a study proposed a VGG-19 model with transfer learning and image segmentation for tomato leaf disease classification [35].

MobileNet is a type of CNN architecture optimized for mobile devices. It is a lightweight and efficient model that can be used for real-time image classification tasks. The study proposed a method for detecting and classifying plant diseases using CNNs based on MobileNet. The authors achieved 98.3% accuracy on a dataset of 38,800 images of 15 different plant diseases [36]. Another study offered a method to identify plant diseases based on leaf patterns using MobileNetV2 [21]. The authors achieved 97.5% accuracy on the same dataset in the previous study. Similarly, Mishra et al. [14] used transfer learning with MobileNet-V1 to identify potato plant lesion features. The authors achieved 94.5% accuracy on a dataset of 1200 potato plant leaf images. A study enhanced deep residual CNNs for plant leaf disease detection using MobileNet, achieving 96.6% accuracy [37]. The authors achieved 96.6% accuracy.

Similar to the previous two architectures, the success of AlexNet had a great impact on the fields of deep learning and computer vision. It demonstrated the potential of deep neural networks for image classification tasks and paved the way for the development of more advanced CNN architectures [38]. Several studies have used the AlexNet architecture to develop deep learning models for agricultural plant disease detection, with satisfactory results [39-41].

ResNet is one of the most popular CNN architectures for image classification. ResNet has proven to be flexible in detecting plant foliar diseases and identifying disease types with high accuracy [42]. A ResNet-based approach has been proposed to detect and classify plant leaf diseases, which are a serious problem for food safety [43]. Another study proposed a customized PDICNet model using ResNet-50 for plant leaf disease identification and classification, employing a deep learning convolutional neural network (DLCNN) classifier model to achieve improved classification performance [44].

Overall, these architectures have proven to be effective deep learning architectures for plant disease detection and classification, with various studies achieving high accuracy in identifying different types of diseases. Their efficiency in processing large amounts of data and classifying them accurately can help in the early detection and prevention of plant diseases, ultimately leading to increased crop yields and economic growth in the agricultural sector.

While there have been many studies focusing on crop-specific disease detection, research specifically addressing potato foliar diseases is relatively limited. Little research has been done to compare the effectiveness of different CNN architectures for this particular application. This research provides a comprehensive comparison of several CNN architectures, namely VGG-16, VGG19, MobileNetV2, AlexNet, and ResNet50 for potato leaf disease classification. This comparison is crucial for advancing the field, as it not only addresses a specific research deficiency but also potentially identifies the most effective deep learning strategies for tackling potato leaf diseases.

The research objective of this study is to provide a comprehensive comparison of various Convolutional Neural Network (CNN) architectures for the classification of potato leaf diseases. This comparison aims to identify the most effective CNN model in accurately detecting and classifying potato leaf diseases, addressing a significant void in the current agricultural research landscape. By evaluating the performance of several leading CNN architectures, including VGG-16, VGG19, MobileNetV2, AlexNet, and ResNet50, this research seeks to contribute valuable insights into the optimization of deep learning techniques for the early detection and prevention of diseases in potato crops, ultimately aiding in the enhancement of agricultural productivity and food security.

3. MATERIALS AND METHODS

3.1 Dataset description

The PlantVillage dataset, an open, accessible, standardized and reliable database created by Hughes & Salathe in 2015 [45], serves as a valuable resource in the fields of agriculture and plant pathology. It comprises a vast collection of annotated images featuring a variety of plant species, including corn, pepper, potato, and tomato, among others. This dataset provides detailed information on the diseases diagnosed for each plant type and is freely available for use in developing and testing various models, including those based on machine learning and deep learning, aimed at disease detection and classification. This research specifically focused on the potato plant diseases section of the PlantVillage dataset, which includes 2,152 potato leaf images categorized into three groups: healthy, early blight, and late blight. Example images of diseased potato plant leaves are displayed in Figure 1. The early and late diseased potato leaf images represent the two stages of potato leaf disease that damage the plant.

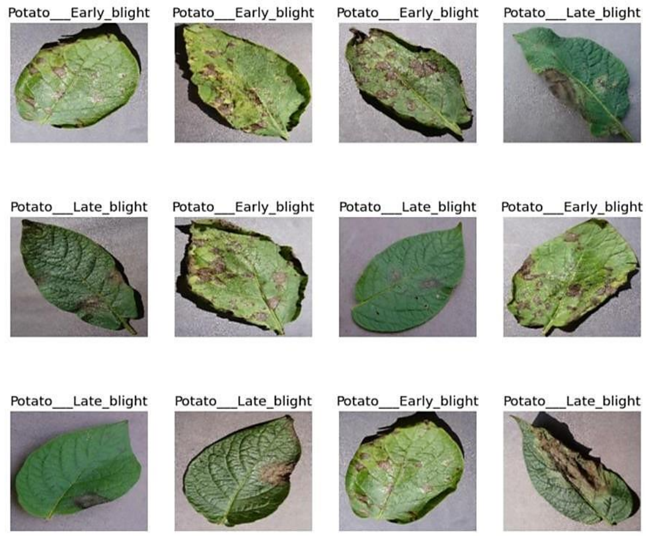

Figure 1. Samples of some diseased potato plant leaves

Table 1 provides data that categorizes leaves into three classes based on their health status or the presence of a particular disease. The identified diseases are Early blight (labeled 0) and Late blight (labeled 1). There is also a category for leaves that are considered Healthy (labeled 2). Both Early blight and Late blight had the same number of samples, 1000 leaves each. The healthy leaf category includes a sample of 152 leaves. Overall, the data consisted of 2,152 samples.

Table 1. Number of potato leaf images in each class of the PlantVillage dataset

Class	Total
Early blight (0)	1,000
Late blight (1)	1,000
Healthy (2)	152
Total	2,152

3.2 Pre-processing

Pre-processing is the process of preparing and refining image data before its utilization in a deep learning model. This stage is crucial in the classification of plant leaf diseases, aiming to enhance the performance and accuracy of the classification model. By effectively preparing image data, pre-processing enables classification models to operate with greater accuracy and efficiency, thereby improving the detection and classification of potato leaf diseases.

In this research, the pre-processing steps undertaken include resizing and scaling. Convolutional Neural Networks (CNNs) require images of uniform size for input. Consequently, all images in the dataset were resized to uniform dimensions of 256×256 pixels. Following resizing, rescaling is applied, wherein the original pixel values ranging from 0 to 255 are scaled down to a range between 0 and 1. This adjustment facilitates model convergence during training and minimizes variations in input values.

The subsequent step involves dividing the dataset into training, validation, and test data, adopting a distribution ratio of 70% for training data, 20% for validation data, and 10% for test data, as detailed in Table 2.

Table 2. Number of images in each training, validation, and testing class

Class	Early Blight	Late Blight	Healthy	Total
Total Data Training	700	700	106	1506
Total Data Validation	200	200	30	430
Total Data Testing	100	100	16	216

The training data are enriched through a data augmentation process to create variations for model development. This process includes generating image variations using techniques such as 45° rotation, along with horizontal and vertical flips. A rotation of 45° is selected for its ability to introduce substantial variation, beneficial for accommodating leaves that may fall or be positioned in a range of unpredictable orientations. By enhancing data variation, data augmentation aids the model in generalizing more effectively to previously unseen data, thereby minimizing the risk of overfitting. Figure 2 shows an example image of the augmentation process.

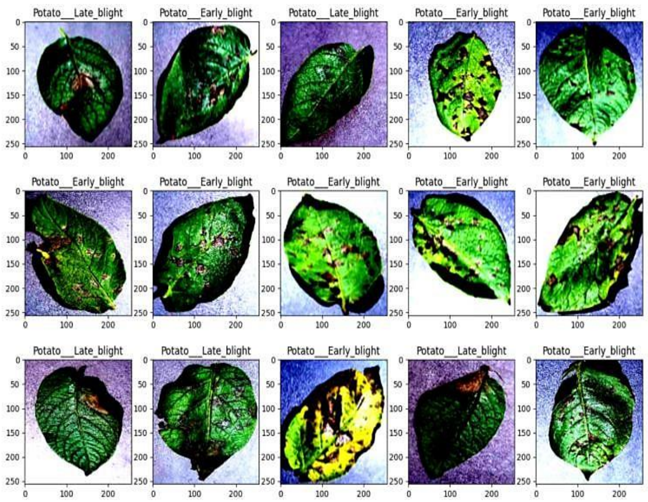

Figure 2. Augmented image

3.3 Model selection and system requirements

Five pre-trained models, namely VGG16, VGG19, MobileNetV2, ResNet50, and AlexNet were explored and trained using the PyTorch library. The selection of these specific architectures was based on their proven superior accuracy over alternative models. Additionally, this choice was guided by a thorough evaluation of their unique features and capabilities, which are particularly suited to the demands of classifying intricate leaf disease patterns. This makes them collectively adept at addressing the complexities and challenges associated with leaf disease imagery.

To aid the processing and analysis of images in this study on potato plant leaf disease detection and classification, additional libraries such as Pillow, NumPy, and Matplotlib were utilized. The integration of these Python libraries into the research workflow markedly enhances this process. Each library serves a distinct purpose, ranging from image manipulation and numerical computations to data visualization, thereby streamlining the research process and facilitating deeper insights into the dataset and the models' performance.

The training process employs Compute Unified Device Architecture (CUDA) technology, which facilitates the efficient use of Graphics Processing Unit (GPU) resources for both preprocessing and training tasks. CUDA harnesses the parallel processing capabilities of GPUs, significantly accelerating not just the model training process but also the preprocessing phase. Deep learning models such as VGG16, VGG19, MobileNetV2, ResNet50, and AlexNet require considerable computational power due to their deep architectures and the complexity of their operations. By utilizing CUDA-enabled GPUs, the time needed for training these models is substantially reduced. This makes it feasible to explore these architectures and hyperparameters within a reasonable timeframe, ensuring a comprehensive application of CUDA's capabilities.

3.4 Hyperparameter tuning

The models were implemented using TensorFlow version 2.13.0 with a total of 2,152 potato leaf images. The models were trained using 1,506 augmented images, with the training spanning 50 epochs and including 5 dropouts. They were then tested using 216 original images. Furthermore, this study used ReLU as the activation function and Stochastic Gradient Descent (SGD) Optimizer with a learning rate of 0.001 to set the weights of the neurons in the neural network, as summarized in Table 3.

Table 3. Summary of model parameters used

Optimizer	Batch Size	Number of Epoch	Dropout	Activation Function	Learning Rate
SGD	32	50	5	ReLU	0,001

3.5 Evaluation

The ternary (3-class) confusion matrix presented in Table 4 shows the classification predictions for three classes: Early blight, Late blight, and Healthy. True Positives (correct predictions) for Early blight, Late blight, and Healthy are abbreviated as 'teb' (true early blight), 'tlb' (true late blight), and 'th' (true healthy), respectively. False Positives (incorrect predictions) include Early blight predicted as Late blight, abbreviated as 'eblb'; Early blight predicted as Healthy is abbreviated as 'ebh'; Late blight predicted as Early blight as 'lbeb'; Late blight predicted as Healthy as 'lbh'; Healthy predicted as Early blight as 'heb'; and Healthy predicted as Late blight as 'hlb'. Correct predictions are represented by the main diagonals (teb, tlb, and th). All non-diagonal elements represent misclassification.

Table 4. Ternary classification confusion matrix

Actual Class	Predicted Class		
	Early blight	Late blight	Healthy
Early blight	teb	eblb	ebh
Late blight	lbeb	tlb	lbh
Healthy	heb	hlb	th

Several measures in computer vision and machine learning are defined based on this classification confusion matrix. For example, the evaluation measures such as precision, recall, F1 Score, accuracy, and specificity, which are used for assessing image classifiers, are shown in Eqs. (1)-(11).

$$Precision_{Earlyblight} = \frac{teb}{teb+lbeb+heb} \quad (1)$$

$$Precision_{Lateblight} = \frac{tlb}{eblb + tlb + hlb} \quad (2)$$

$$Precision_{Healthy} = \frac{th}{ebh + lbh + th} \quad (3)$$

$$Recall_{Earlyblight} = \frac{teb}{teb + eblb + ebh} \quad (4)$$

$$Recall_{Lateblight} = \frac{tlb}{tlb + lbeb + lbh} \quad (5)$$

$$Recall_{Healthy} = \frac{th}{th+heb+hlb} \quad (6)$$

$$Specificity_{Earlyblight} = \frac{tlb + th}{tlb + th + lbeb + heb} \quad (7)$$

$$Specificity_{Lateblight} = \frac{teb + th}{teb + th + eblb + hlb} \quad (8)$$

$$Specificity_{Healthy} = \frac{teb + tlb}{teb + tlb + ebh + lbh} \quad (9)$$

$$F1Score = \frac{2(Recall.Precision)}{(Recall + Precision)} \quad (10)$$

$$Accuracy = \frac{teb + tlb + th}{teb + tlb + th + feb + flb + fh} \quad (11)$$

4. RESULTS AND DISCUSSION

Table 5 shows data for 216 samples divided into 3 classifications using the VGG16 architecture. Of these, 204 samples (94.4%) were correctly classified, leaving 12 samples (5.6%) misclassified. In the Early blight category, 94 samples were correctly predicted, while 5 were incorrectly predicted as Late blight and one as Healthy, totaling 6 misclassifications. The Late blight category had 96 correct predictions, with 3 misclassified as Early blight and 1 as Healthy, for a total of 4 misclassifications. Finally, in the Healthy category, 14 samples were classified correctly, but 2 were misclassified: one as Early blight and the other as Late blight.

Table 5. Confusion matrix of classification results

Actual Class	Predicted Class		
	Early blight	**Late blight**	**Healthy**
Early blight	94	5	1
Late blight	3	96	1
Healthy	1	1	14

The classification report and confusion matrix for the five architectures is shown in Table 6. Both tools provide valuable insights and together can provide a comprehensive view of the model's performance across all classes. All architectures performed well across all three classes, with classes 0 and 1 having the highest number of instances. Class 2, representing the Healthy category, has fewer instances and slightly lower precision, recall, and F1Score compared to the other two classes.

The classification report and confusion matrix reveal that VGG16 and VGG19 deliver high performance in precision and recall, signifying their capability to accurately identify potato leaf diseases with minimal error. However, their extensive depth necessitates greater computational resources, which should be factored into their practical deployment. A high support value for certain classes suggests adequate training data for those specific conditions, but it may also reflect a potential bias towards more frequently occurring classes in the dataset.

The classification report for MobileNetV2 demonstrates a well-balanced trade-off between precision and recall, with its f1-score evidencing a robust capability to classify various disease classes. Uniform support across classes implies that the model has been trained on a balanced dataset, which is advantageous for its generalization capabilities. Designed with mobile applications in mind, MobileNetV2 has faster inference times compared to VGG16 and VGG19, making it highly beneficial for real-time applications.

ResNet50’s classification report indicates the most consistent performance among the models, with high precision, recall, and f1-score values, reflecting its exceptional ability to recognize diverse disease classes. This consistency is attributed to ResNet50's use of skip connections, which effectively address the vanishing gradient problem common in deep networks such as VGG16 and VGG19.

As one of the pioneering architectures, AlexNet does not have the architectural complexity of newer models, but still shows good performance. The classification report suggests that AlexNet remains a competitive choice, particularly when limited by resources, although it may not be as efficient as MobileNetV2 or as accurate as ResNet50.

Analysis of the classification reports and confusion matrices provides insight into each model's performance under similar conditions. By comparing the results of these reports on different architectures for detecting potato leaf diseases, researchers or practitioners can determine which models are most suitable for implementation in real agricultural settings based on criteria such as accuracy, speed of inference, and resource requirements.

Table 7 displays the performance metrics of five popular CNN architectures: VGG16, VGG19, MobileNetV2, ResNet50, and AlexNet, including training accuracy, validation accuracy, training loss, validation loss, and testing accuracy. Training and validation accuracies were relatively consistent across models, ranging from 92% to 96% for training accuracy and 93% to 96% for validation accuracy. This consistency indicates that all models have learned reasonably well from the training data and have a comparable ability to generalize to unseen data. The training and validation loss values are also consistent across models. The values indicate that none of the models significantly overfits the training data, as the losses are relatively close for both the training and validation sets. The testing accuracy, which measures the performance of the architecture on unseen data, shows little variation. ResNet50 achieved the highest testing accuracy at 97%, followed by MobileNetV2 and AlexNet, both at 95%. VGG16 and VGG19 achieved 94% and 90% accuracy, respectively. ResNet50 showed the best overall performance for this particular task, as it had the highest testing accuracy. All pre-trained models showed relatively good generalization as validation and testing accuracies were close. ResNet50 stands out as an architecture with relatively low training and validation losses, indicating that it can be computationally efficient while maintaining strong performance. ResNet50 appears to be the best in terms of testing accuracy and training efficiency on this dataset.

Table 6. Classification report and confusion matrix of five different architectures, (a) VGG16, (b) VGG19, (c) MobileNetV2, (d) ResNet50, (e) AlexNet

```
              precision    recall  f1-score   support

           0       0.96      0.94      0.95       100
           1       0.94      0.96      0.95       100
           2       0.88      0.88      0.88        16

    accuracy                           0.94       216
   macro avg       0.93      0.92      0.92       216
weighted avg       0.94      0.94      0.94       216
```

Actual Class	Predicted Class		
	Early Blight	**Late Blight**	**Healthy**
Early blight	94	5	1
Late blight	3	96	1
Healthy	1	1	14

(a) Classification report and confusion matrix of VGG16

```
              precision    recall  f1-score   support

           0       0.94      0.90      0.92       100
           1       0.92      0.92      0.92       100
           2       0.60      0.75      0.67        16

    accuracy                           0.90       216
   macro avg       0.82      0.86      0.84       216
weighted avg       0.90      0.90      0.90       216
```

Actual Class	Predicted Class		
	Early Blight	**Late Blight**	**Healthy**
Early blight	90	6	4
Late blight	4	92	4
Healthy	2	2	12

(b) Classification report and confusion matrix of VGG19

```
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       100
           1       0.95      0.96      0.96       100
           2       0.82      0.88      0.85        16

    accuracy                           0.95       216
   macro avg       0.92      0.93      0.92       216
weighted avg       0.95      0.95      0.95       216
```

Actual Class	Predicted Class		
	Early Blight	**Late Blight**	**Healthy**
Early blight	96	4	0
Late blight	1	96	3
Healthy	1	1	14

(c) Classification report and confusion matrix of MobileNetV2

```
              precision    recall  f1-score   support

           0       0.96      0.97      0.97       100
           1       0.98      0.98      0.98       100
           2       0.93      0.88      0.90        16

    accuracy                           0.97       216
   macro avg       0.96      0.94      0.95       216
weighted avg       0.97      0.97      0.97       216
```

Actual Class	Predicted Class		
	Early Blight	**Late Blight**	**Healthy**
Early blight	97	2	1
Late blight	2	98	0
Healthy	2	0	14

(d) Classification report and confusion matrix of ResNet50

```
              precision    recall  f1-score   support

           0       0.96      0.94      0.95       100
           1       0.96      0.96      0.96       100
           2       0.83      0.94      0.88        16

    accuracy                           0.95       216
   macro avg       0.92      0.95      0.93       216
weighted avg       0.95      0.95      0.95       216
```

Actual Class	Predicted Class		
	Early Blight	**Late Blight**	**Healthy**
Early blight	94	4	2
Late blight	3	96	1
Healthy	1	0	15

(e) Classification report and confusion matrix of AlexNet

Table 7. Summary of training, validation, and testing results using five pre-trained CNN

Pre-Trained Model	Training Acc (%)	Validation Acc (%)	Training Loss	Validation Loss	Testing Acc (%)
VGG16	95	94	0.0358	0.0430	94
VGG19	92	94	0.2935	0.0555	90
MobileNetV2	94	95	0.0771	0.0630	95
ResNet50	96	96	0.0174	0.0171	97
AlexNet	94	93	0.0602	0.0966	95

Table 8 presents predictions from a classification task aimed at identifying diseases across three potato leaf classes. VGG16 performed well in all classes, with high precision, recall, and F1Score for Early blight and Late blight the diseases. However, VGG16 showed lower performance for the Healthy class, with accuracy, precision, recall, and an F1Score of 88%. The negative prediction rate was highest for the Healthy class, indicating a higher rate of misclassification within this category.

Table 8. Comparison of architectural performance in measuring accuracy, precision, recall, and f1 score

Architectures Performance	Classes		
	Early Blight	Late Blight	Healthy
VGG16			
Accuracy	96%	94%	88%
Precision	96%	94%	88%
Recall	94%	96%	88%
F1 Score	95%	95%	88%
Positive Prediction	94%	96%	88%
Negative Prediction	6%	4%	12%
VGG19			
Accuracy	94%	92%	60%
Precision	94%	92%	60%
Recall	90%	92%	75%
F1 Score	92%	92%	67%
Positive Prediction	90%	92%	75%
Negative Prediction	10%	8%	25%
MobileNetV2			
Accuracy	98%	95%	82%
Precision	98%	95%	82%
Recall	96%	96%	88%
F1 Score	97%	96%	85%
Positive Prediction	96%	94%	88%
Negative Prediction	4%	4%	12%
ResNet50			
Accuracy	96%	96%	83%
Precision	96%	98%	93%
Recall	97%	98%	88%
F1 Score	97%	98%	90%
Positive Prediction	97%	98%	88%
Negative Prediction	3%	2%	12%
AlexNet			
Accuracy	96%	97%	82%
Precision	96%	96%	83%
Recall	94%	96%	94%
F1 Score	95%	96%	88%
Positive Prediction	94%	96%	94%
Negative Prediction	6%	4%	6%

VGG19 also performed relatively well for Early blight and Late blight, with high precision, recall, and F1 scores for both classes. However, it performed significantly lower for the Healthy class, with 60% accuracy. This architecture had the highest recall (75%) for the Healthy class but also exhibited a higher rate of false negatives prediction for the Healthy class. MobileNetV2 stands out for its outstanding performance in classifying Early blight and Late blight, with high accuracy, precision, recall, and F1 scores. However, it is less accurate for the Healthy class, with accuracy, precision, recall, and F1 scores of 82%, 82%, 88%, and 85%, respectively.

ResNet50 excelled in classifying Late blight, with high accuracy, precision, recall, and F1 score. ResNet50 also performed well for Early blight, with high accuracy and balanced precision and recall. However, it performed relatively lower for the Healthy class, with 83% accuracy. AlexNet performed well for Late blight and Early blight, with high accuracy, precision, recall, and F1 scores. It performed slightly lower for the Healthy class, with an accuracy of 82%. The architecture had a relatively low negative prediction rate across all classes, indicating a lower level of misclassification than other classes.

Each architecture has strengths and weaknesses in classifying the three classes, with varying levels of accuracy, precision, recall, and F1 score. The Healthy class consistently presented challenges across all architectures, with a higher negative prediction rate, indicating a tendency to misclassify non-healthy instances as Healthy.

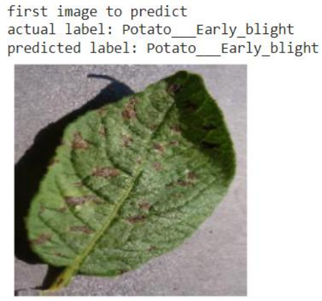

Figure 3. First image to predict

The prediction for the first image was as follows: the actual label is Potato Early blight, and the predicted label is also Potato Early blight. The model correctly predicted the class for the first image as Potato Early blight, which matches the actual label (Figure 3).

The accuracy and specificity values of the different CNN architectures when applied with transfer learning are shown in Table 9. Specificity is an important metric in many contexts, including image classification. Specificity for each class is calculated by considering one class as the positive class and the rest as the negative class. It provides insight into how well the classification system identifies negative classes. This is particularly important in cases where avoiding false positives is as critical as identifying true positives. In datasets where one class significantly outnumbers the others (imbalanced datasets), like in this study where the healthy leaf dataset is not balanced with the other two, specificity becomes an important measure. It ensures that the model's ability to correctly identify the less frequent class (often the negative class) is not overlooked. ResNet50 has the highest accuracy (97%) among the given architectures. ResNet50 also achieved the highest specificity (98%). On the other hand, VGG19 has both the lowest accuracy (90%) and the lowest specificity (94%).

Table 9. Accuracy and specificity score of different CNN architectures with transfer learning

Architecture(s)	Accuracy (%)	Specificity (%)
VGG16	94	97
VGG19	90	94
MobileNetV2	95	97
ResNet50	97	98
AlexNet	95	97

A snapshot of the prediction results for three classes of foliar diseases of potato plants is shown in Figure 4. The data presented highlight the predictive model’s performance in classifying potato leaves into three health status categories: Early blight, Late blight, and Healthy. The figure demonstrates remarkable accuracy in these predictions. In each case, the predicted label matches the actual label. The model exhibits high confidence in its predictions, with confidence scores often approaching 100%. The lowest confidence score was 98.68%, which is still very high in the context of a classification task.

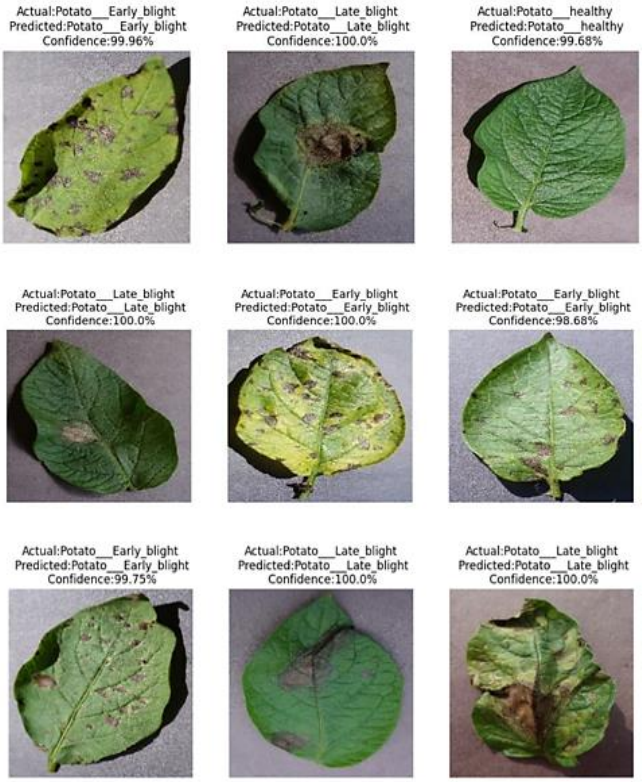

Figure 4. Image prediction results

On multiple instances of the same category, the model consistently predicted the correct label. This shows that the model is not only accurate but also reliable in its predictions for this dataset. Overall, the model is very good at classifying potato leaves based on their health status. The model not only provides correct classification but also shows high confidence in its predictions. This high performance indicates that the model has been well-trained and most likely exposed to a wide variety of samples during the training phase. The results presented make a compelling case for using the model in real-world applications, given its demonstrated ability to identify the health status of potato leaves with remarkable accuracy and confidence.

Table 10 provides a comprehensive review of research studies on potato leaf disease detection, using a variety of computational models from 2019 to 2024. These studies employ a range of methods, predominantly Convolutional Neural Network (CNN) architectures, in addition to traditional machine learning techniques.

Afonso et al. [7] utilized ResNet18 and achieved a notable 95% accuracy. The table also details various approaches such as the InceptionV3 model by Chugh et al. [9], which yielded a 90% accuracy, and a modified AlexNet architecture by Bajpai et al. [10] that reached a 61% accuracy.

The R-CNN-GC model, employed by SERT [11], achieved 94% accuracy. Nazir et al. [12] attained a 98% accuracy with EfficientPNet, and Mishra et al. [14] led with a 99% accuracy using MobileNet-V1.

Furthermore, the table highlights the ongoing challenge in accurately classifying healthy leaves, as evidenced by the varying performance of models such as VGG16 and VGG19 architectures by Sholihati et al. [16], which achieved accuracies ranging from 91% to 93%.

Additionally, the table lists more traditional machine learning methods such as Random Forest, Logistic Regression, K-Nearest Neighbors, Decision Tree, Naïve Bayes, Linear Discriminant Analysis, and Support Vector Machine. These methods have achieved accuracies ranging from 91% to 97%, as implemented by Iqbal and Tadukler [25].

Table 10. Comparison of the proposed model with state-of-the-art models in potato leaf disease detection accuracy

Author(s) / References	Year	Method	Accuracy (%)
Afonso et al. [7]	2019	ResNet18	95
Lee et al. [8]	2020	CNN	99
Chugh, et al. [9]	2020	InceptionV3	90
Bajpai, et al. [10]	2023	Modified AlexNet	61
SERT [11]	2021	R-CNN-GC	94
Nazir et al. [12]	2023	EfficientPNet	98
Mishra et al. [14]	2021	MobileNet-V1	99
Sholihati et al. [16]	2020	VGG16 and VGG19	91-93
Hylmi at al. [22]	2022	Support Vector Machine	97
		VGG16	97
Nishad et al. [23]	2022	VGG19	94
		ReNet50	67
Rusli et al. [24]	2022	K-Means	94
		Random Forest	97
		Logistic Regression	94
		K-Nearest Neighbors	91
Iqbal and Talukder [25]	2020	Decision Tree	91
		Naïve Bayes	84
		Linear Discriminant Analysis	78
		Support Vector Machine	37
		VGG16	94
		VGG19	90
Proposed Model	**2024**	MobileNetV2	95
		ResNet50	**97**
		AlexNet	95

This research presents a comparative study of a multi-model approach that includes VGG16, VGG19, MobileNetV2, ResNet50, and AlexNet, with respective accuracies of 94%, 90%, 95%, 97%, and 95%. This study posits the proposed models as competitive counterparts to existing state-of-the-art models, potentially providing novel insights into potato leaf disease detection and advancing agricultural technology.

5. CONCLUSION AND FUTURE WORK

The evaluation of various pre-trained convolutional neural network architectures: VGG16, VGG19, MobileNetV2, ResNet50, and AlexNet, revealed different strengths and weaknesses in classifying the three leaf conditions. Notably, ResNet50 emerged as the most outstanding, achieving the highest testing accuracy of 97% and specificity of 98%. Conversely, VGG19 had the lowest values for accuracy and specificity, at 90% and 94% respectively. Although several architectures showed commendable performance in classifying potato leaf diseases, ResNet50 proved to be the most effective and efficient model in this study. By providing reliable and efficient disease detection methodologies, such research can significantly contribute to sustainable farming practices, reducing pesticide use, and enhancing crop management strategies globally. This research highlights the importance of specificity as a metric, especially in contexts such as image classification, to determine accuracy in identifying negative or non-diseased classes. For future research, exploring the integration of these CNN models with IoT-based agricultural sensors could not only open new avenues for real-time field monitoring and disease management but also enhance precision agriculture practices. Additionally, investigating the adaptability of these models to other crop diseases and environmental conditions would further validate their utility in diverse agricultural scenarios, paving the way for more resilient food production systems.

ACKNOWLEDGMENT

We sincerely thank the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia for its unwavering support and commitment as a research funder under Contract No. 0536/E5/PG.02.00/2023.

REFERENCES

[1] Akinyi, D.P., Ng'ang'a, S.K., Ngigi, M., Mathenge, M., Girvetz, E. (2022). Cost-benefit analysis of prioritized climate-smart agricultural practices among smallholder farmers: evidence from selected value chains across sub-Saharan Africa. Heliyon, 8(4): 1-11. https://doi.org/10.1016/j.heliyon.2022.e09228

[2] Burgos, G., Zum Felde, T., Andre, C., Kubow, S. (2020). The potato and its contribution to the human diet and health. The Potato Crop, 37-74. https://doi.org/10.1007/978-94-011-2340-2

[3] Rupp, J., Jacobsen, B. (2017). Bacterial and fungal diseases of potato and their management. Montana State University, January, 1-12.

[4] Boulent, J., Foucher, S., Théau, J., St-Charles, P.L. (2019). Convolutional neural networks for the automatic identification of plant diseases. Front. Plant Sci., 10: 1-15. https://doi.org/10.3389/fpls.2019.00941

[5] Hassan, S.M., Maji, A.K., Jasiński, M., Leonowicz, Z., Jasińska, E. (2021). Identification of plant-leaf diseases using CNN and transfer-learning approach. Electronics, 10(12): 1-19. https://doi.org/10.3390/electronics10121388

[6] Karunanithi, A., Singh, A.S., Kannapiran, T. (2022). Enhanced hybrid neural networks (CoAtNet) for paddy crops disease detection and classification. Revue d'Intelligence Artificielle, 36(5): 671-679. https://doi.org/10.18280/ria.360503

[7] Afonso, M., Blok, P.M., Polder, G., Van Der Wolf, J.M., Kamp, J. (2019). Blackleg detection in potato plants using convolutional neural networks. IFAC-PapersOnLine, 52(30): 6-11. https://doi.org/10.1016/j.ifacol.2019.12.481

[8] Lee, T.Y., Yu, J.Y., Chang, Y.C., Yang, J.M. (2020). Health detection for potato leaf with convolutional neural network. In Indo - Taiwan 2nd International Conference on Computing, Analytics and Networks, Indo-Taiwan ICAN 2020, Rajpura, India, pp. 289-293. https://doi.org/10.1109/Indo-TaiwanICAN48429.2020.9181312

[9] Chugh, G., Sharma, A., Choudhary, P., Khanna, R. (2020). Potato leaf disease detection using inception V3. International Research Journal of Engineering and Technology (IRJET), 7(11): 1363-1366.

[10] Bajpai, A., Tyagi, M., Khare, M., Singh, A. (2023). A robust and accurate potato leaf disease detection system using modified AlexNet model. In 9th International Conference on Computer and Communication Engineering, Kuala Lumpur, Malaysia, pp. 264-269. https://doi.org/10.1109/ICCCE58854.2023.10246064

[11] SERT, E. (2023). A deep learning based approach for the detection of diseases in pepper and potato leaves. Anadolu Anadolu Journal of Agricultural Sciences, 36: 167-178. https://doi.org/10.7161/omuanajas.805152

[12] Nazir, T., Iqbal, M.M., Jabbar, S., Hussain, A., Albathan, M. (2023). EfficientPNet—an optimized and efficient deep learning approach for classifying disease of potato plant leaves. Agriculture, 13(4): 1-18. https://doi.org/10.3390/agriculture13040841

[13] Nandhini, S., Ashokkumar, K. (2022). An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. Neural Computing and Applications, 34(7): 5513-5534. https://doi.org/10.1007/s00521-021-06714-z

[14] Mishra, S., Singh, A., Singh, V. (2021). Application of MobileNet-v1 for potato plant disease detection using transfer learning. ACM International Conference Proceeding Series, Workshop on Algorithm and Big Data, pp. 14-19. https://doi.org/10.1145/3456389.3456403

[15] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 1-12. https://doi.org/10.1109/CVPR.2016.90

[16] Sholihati, R.A., Sulistijono, I.A., Risnumawan, A., Kusumawati, E. (2020). Potato leaf disease classification using deep learning approach. In IES 2020 - International Electronics Symposium: The Role of Autonomous and Intelligent Systems for Human Life and Comfort,

Surabaya, Indonesia, pp. 392-397. https://doi.org/10.1109/IES50839.2020.9231784
[17] Varga, D. (2020). Multi-pooled inception features for no-reference image quality assessment. Applied Sciences, 10(6): 1-15. https://doi.org/10.3390/app10062186
[18] Hidayatuloh, A., Nursalman, M., Nugraha, E. (2018). Identification of tomato plant diseases by leaf image using Squeezenet model. International Conference on Information Technology Systems and Innovation, ICITSI, Bandung, Indonesia, pp. 199-204. https://doi.org/10.1109/ICITSI.2018.8696087
[19] Tan, M., Le, Q.V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. https://doi.org/10.48550/arXiv.1905.11946
[20] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q. (2017). Densely connected convolutional networks. In 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 2261-2269. https://doi.org/10.1109/CVPR.2017.243
[21] Verma, D., Bordoloi, D., Tripathi, V. (2021). Plant leaf disease detection using mobilenetv2. Webology, 18(5): 3241-3246. https://doi.org/10.29121/web/v18i5/60
[22] Hylmi, M.S., Wiharto, Suryani, E. (2022). Detection of potato leaf disease using multi-class support vector machine based on texture, color, and shape features. In International Conference on Electrical and Information Technology, Malang, Indonesia, pp. 20-24. https://doi.org/10.1109/IEIT56384.2022.9967866
[23] Nishad, M.A.R., Mitu, M.A., Jahan, N. (2022). Predicting and classifying potato leaf disease using K-means segmentation techniques and deep learning networks. Procedia Computer Science, 212: 220-229. https://doi.org/10.1016/j.procs.2022.11.006
[24] Rusli, A.H.T., Meng, B.C.C., Damanhuri, N.S., Othman, N.A., Othman, M.H., Zaidi, W.F.A.W. (2022). Potato leaf disease classification using image processing and artificial neural network. In 12th IEEE International Conference on Control System, Computing and Engineering, Penang, Malaysia, pp. 107-112. https://doi.org/10.1109/ICCSCE54767.2022.9935654
[25] Iqbal, M.A., Talukder, K.H. (2020). Detection of potato disease using image segmentation and machine learning. In International Conference on Wireless Communications, Signal Processing and Networking, Chennai, India, pp. 43-47. https://doi.org/10.1109/WiSPNET48689.2020.9198563
[26] Sarker, I.H. (2021). Machine Learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3): 1-21. https://doi.org/10.1007/s42979-021-00592-x
[27] Saleem, M.H., Potgieter, J., Arif, K.M. (2020). Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers. Plants, 9(10): 1-17. https://doi.org/10.3390/plants9101319
[28] Jackulin, C., Murugavalli, S. (2022). A comprehensive review on detection of plant disease using machine learning and deep learning approaches. Measurement: Sensors, 24: 1-10. https://doi.org/10.1016/j.measen.2022.100441
[29] Lu, J., Tan, L., Jiang, H. (2021). Review on convolutional neural network (CNN) applied to plant leaf disease classification. Agriculture, 11(8): 707. https://doi.org/10.3390/agriculture11080707
[30] Babu, P.R., Krishna, A.S. (2023). Deep learning-assisted SVMs for efficacious diagnosis of tomato leaf diseases: A comparative study of GoogleNet, AlexNet, and ResNet-50. Ingénierie des Systèmes d'Information, 28(3): 639-645. https://doi.org/10.18280/isi.280312
[31] Wagle, S.A., Harikrishnan, R. (2021). Comparison of plant leaf classification using modified AlexNet and support vector machine. Traitement du Signal, 38(1): 79-87. https://doi.org/10.18280/TS.380108
[32] Kumar, A., Kumar, A. (2023). Plant Disease Detection using VGG16. International Journal of Creative Research Thoughts (IJCRT), 11(1): 614-619. https://ijcrt.org/papers/IJCRT2301347.pdf.
[33] Alatawi, A.A., Alomani, S.M., Alhawiti, N.I., Ayaz, M. (2022). Plant disease detection using AI based VGG-16 model. International Journal of Advanced Computer Science and Applications (IJACSA), 13(4): 718-727. https://doi.org/10.14569/IJACSA.2022.0130484
[34] Paymode, A.S., Malode, V.B. (2022). Transfer learning for multi-crop leaf disease image classification using Convolutional Neural Network VGG. Artificial Intelligence in Agriculture, 6: 23-33. https://doi.org/10.1016/j.aiia.2021.12.002
[35] Nguyen, T.H., Nguyen, T.N., Ngo, B.V. (2022). A VGG-19 model with transfer learning and image segmentation for classification of tomato leaf disease. AgriEngineering, 4(4): 871-887. https://doi.org/10.3390/agriengineering4040056
[36] Ashwinkumar, S., Rajagopal, S., Manimaran, V., Jegajothi, B. (2021). Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks. Materials Today: Proceedings, 2021, 51(1): 480-487. https://doi.org/10.1016/j.matpr.2021.05.584
[37] Arun, P.J., Kanchanadevi, K., Rajalakshmi, N.R., Arulkumaran, G. (2022). An Improved deep residual convolutional neural network for plant leaf disease detection. Computational Intelligence and Neuroscience, 34(28): 1-9. https://doi.org/10.1002/cpe.7357
[38] Soujanya, K., Jabez, J. (2021). Recognition of plant diseases by leaf image classification based on improved AlexNet. In 2nd International Conference on Smart Electronics and Communication, Trichy, India, pp. 1306-1313. https://doi.org/10.1109/ICOSEC51865.2021.9591809
[39] Chen, H.C., et al. (2022). AlexNet convolutional neural network for disease detection and classification of tomato leaf. Electronics, 11(6): 1-17. https://doi.org/10.3390/electronics11060951
[40] Yeh, J.F., Wang, S.Y., Chen, Y.P. (2021). Crop disease detection by image processing using modified Alexnet. In 3rd IEEE Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability, ECBIOS, Tainan, Taiwan, pp. 159-165. https://doi.org/10.1109/ECBIOS51820.2021.9510426
[41] Li, Z., Li, C., Deng, L.F., et al. (2022). Improved AlexNet with inception-V4 for plant disease diagnosis. Computational Intelligence and Neuroscience, 2022: 1-12. https://doi.org/10.1155/2022/5862600
[42] Loganathan, P., Karthikeyan, R., Scholar, R. (2021). Residual neural network (ResNet) based plant leaf disease detection and classification. Turkish Online Journal of Qualitative Inquiry (TOJQI), 12(6): 1395-1401.

[43] Kumar, V., Arora, H., Harsh, Sisodia, J. (2020). ResNet-based approach for detection and classification of plant leaf diseases. In International Conference on Electronics and Sustainable Communication Systems, ICESC, Coimbatore, India, pp. 495-502. https://doi.org/10.1109/ICESC48915.2020.9155585

[44] Reddy, S.R.G., Varma, G.P.S., Davuluri, R.L., (2023) Resnet-based modified red deer optimization with DLCNN classifier for plant disease identification and classification. Computers and Electrical Engineering, 105: 1-15. https://doi.org/10.1016/j.compeleceng.2022.108492

[45] Hughes, D.P., Salathe, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv:1511.08060.