LeafVision: An Improved ResNet+U-Net based Deep Learning Framework for Potato Leaf Disease Detection

Sarowar Morshed Shawon1, a), Fahima Lokman Niha2, b), Steve Austin2, c), Shah Nawaz Haider2, d), Arnab Barua1, e), Jotika Das2, f) and H.T. Zubair3, 4, g)

1Department of Electrical and Electronic Engineering, University of Science and Technology Chittagong, Chattogram 4202, Bangladesh

2Department of Computer Science and Engineering, University of Science and Technology Chittagong, Chattogram 4202, Bangladesh

3Center for Fiber Networking and Communication, COE for Intelligent Network, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia.

4Faculty of Computing and Informatics (FCI), Multimedia University, 63100 Cyberjaya, Selangor, Malaysia.

g) *Corresponding author:* [*zubair@mmu.edu.my*](mailto:zubair@mmu.edu.my)a)[*shawon@ustc.ac.bd*](mailto:shawon@ustc.ac.bd)

b)[*nihafahima9@gmail.com*](mailto:nihafahima9@gmail.com)

c)[*steve\_austin\_@outlook.com*](mailto:steve_austin_@outlook.com)

d)[*nawazhaider60@gmail.com*](mailto:nawazhaider60@gmail.com)

e)[a*rnab.me@outlook.com*](mailto:arnab.me@outlook.com)

f)[*jotikadas57@gmail.com*](mailto:jotikadas57@gmail.com)

**Abstract.** Global food security is seriously threatened by potato leaf diseases, which call for early and precise identification to reduce crop losses. The rapid growth of AI technology and its adoption in agriculture makes it imperative to identify plant diseases at an earlier stage for sustainable crop production. Some diseases can be truly harmful to potato crops as they reduce both productivity and quality. To mitigate them, AI-driven technologies are very crucial since manual diagnosis takes a lot of time and requires specialized knowledge. In this paper, we proposed a Deep Learning (DL) approach to improve accuracy by fine-grained classification of Potato Leaf Diseases (PLD) into seven groups: Bacteria, Fungi, Phytophthora, Nematode, Pests, Virus and Healthy. The proposed improved ResNet+U-Net based approach, exceeds the traditional standalone and hybrid models and achieves 97.9% accuracy and 98% precision, demonstrating its high effectiveness and scalability in PLD detection. Implementation of advanced segmentation approaches enhances the understanding of disease progression and enables more precise and effective treatment for better crop management. The LeafVision model was tested and evaluated against the PLD dataset in uncontrolled environment, to assess its performance in uncontrolled environmental conditions and benchmarked against leading methods to identify improvements in accuracy, computational efficiency, or both. The proposed technique proved superior to recreated benchmark models like VGG and standalone ResNet50, achieving remarkable improvements in its evaluation criteria.

# Introduction

Potato (*Solanum tuberosum*) is an essential and widely cultivated crop that belongs to the Solanaceae family. It is a starchy tuberous crop that serves as a staple food for millions of people around the world [1]. Selection of the perfect and resilient variety, hence, is the most crucial step of cultivating potato and recent advances in remote sensing of environmental data, using IoT, has vital contributions to profiling [2] and maximum production [3]. However, potato crops face serious threats from various diseases which cause massive economic losses and lowers agriculture productivity. So, early detection and accurate diagnosis of these diseases are essential for effective disease management and control [4]. To effectively manage these issues and minimize crop losses, farmers and local experts typically rely on visual inspection [5]. Traditional methods of disease detection are very time consuming and highly inaccurate. Potato leaf disease detection, in its early stage is challenging because of variations in crop species, crop diseases symptoms and environmental factors [6]. There are several diseases we can see in potato leaf such as bacteria which causes both wilting and yellowing of the leaves, fungi which lead to lesions and spots on leaves, pests like aphid’s damage leaves, virus causes the mosaic patterns and discoloration, nematodes damage the leaves and roots, phytophthora causes rapid decay on potato leaves. In recent years, machine learning and deep learning has become very popular in agriculture because it can extract complex pattern and features from large data-sets and its capabilities to contribute to the crop yield [7] through recommendation of crops [8] and its yield prediction [9]. In plant disease detection, image segmentation plays an important role by precisely identifying diseased areas within the leaves, rather than simply classifying the entire image. This study focuses on identifying seven categories of potato leaves, six representing diseases (Bacteria, Fungi, Phytophthora, Nematode, Pests, and Virus) and one representing healthy leaves. This study present LeafVision, an enhanced ResNet+U-Net DL model designed for classification using segmented images. Unlike previous works, our approach uses data collected from uncontrolled environments, adding complexity to the task. The method aims to enable early and accurate disease detection, supporting precision agriculture by helping farmers improve crop yield, reduce losses and manage potato leaf diseases more efficiently.

# Literature Survey

The relevant studies, in 2021 Rashid et al. [6] combine YOLOv5 with a custom CNN architecture to detect early blight, late blight and healthy. In this study, dataset was collected from Punjab, Pakistan and plant village. The multi-level model achieved 99.75% accuracy. In this study authors demonstrate robustness with data augmentation techniques. In 2023, Acharjee et al. [10] uses several deep learning models like CNN, SVM, and VGG-16 for detecting between healthy and unhealthy leaves. In this study, CNN model shows highest accuracy of 98%. Iqbal and Talukder in 2020 [11] studied how to identify and categorize potato leaf diseases using image segmentation methods. The authors used 450 images of healthy and diseased potato leaves from plant village dataset. To find the different between healthy and diseased leaves, they used seven classifier algorithm, and random forest shows the best accuracy 97%. Bonik et al. in 2023 [12] developed a sequential CNN model to detect healthy, early blight, and late blight potato leaves. In this study authors used 3561 images from plant village and field dataset. The sequential CNN model achieved 94.20% accuracy.

Chen et al. in 2023 [13] describe a lightweight model named MobS\_Net for weekly supervised learning in potato disease detection. The authors used plant village dataset of 3276 images. The MobS\_Net model achieved 97.73% accuracy under a noisy and real-world scenario. In 2021, Saeed et al. [14] used advanced deep learning models like ResNet-152 and InceptionV3 for classify three types of potato leaf disease such as healthy, early blight, late blight. The authors used 1500 images from Kaggle and ResNet-152 achieve highest accuracy 98.34%. Adluri et al.1 in 2024 explored the use of VGG16 a deep learning model for detecting potato leaf diseases such as late blight, early blight and anthracnose and achieved 97% classification accuracy. In 2024, Rdewan et al. [15] investigated the classification of potato leaf diseases using some optimized machine learning models like SVM, Random Forest and KNN. The dataset has 4020 records with powerful weather information such as temperature, humidity, wind speed, wind direction, visibility, and atmospheric pressure. A multilayer perceptron (MLP) model achieved the highest accuracy of 98.30%.

Despite the remarkable progress made in potato leaf disease detection using various machine learning and deep learning models such as YOLOv5, CNNs, VGG-16, ResNet and ensemble classifiers like Random Forest and SVM, most existing approaches primarily focus on classification rather than precise localization or segmentation of disease-affected regions. While models like YOLOv5 and CNN architectures have achieved high classification accuracies (ranging from 94.2% to 99.75%), they lack pixel-level disease region identification, which is crucial for accurate disease severity assessment and targeted intervention. Furthermore, current studies often rely on limited datasets, encompassing less variety of diseases, or do not generalize well under real-world conditions due to the absence of varied data and spatial feature mapping. This creates a significant research gap in implementing segmentation-based architectures tailored to leaf disease detection. To address this, this study proposes LeafVision, an improved ResNet+U-Net based deep learning framework capable of not only classifying but also segmenting infected areas on potato leaves. The encoding capabilities of ResNet50, capable of extracting deep, hierarchical feature and overcome vanishing gradient problem by skipping residual connections; while the decoder architecture of U-Net, known for its strength in biomedical image segmentation, is leveraged to enable precise, end-to-end disease localization thus advancing the state-of-the-art in plant pathology automation.

# Methods

## Data Collection and Preprocessing

The dataset collected had been taken from Mendeley data with the titled “Potato Leaf Disease Dataset in Uncontrolled Environment” [16]. It contains 7 classes of which 6 are diverse diseases which are Bacteria, Fungi, Nematode, Pest, Phytopthora and Virus, and the class of healthy leaves. This dataset had a total of 3076 pictures and the data was captured in an uncontrolled setting which results in a wide range of variables, which include background, direction and distance. The 6 classes of disease have been shown in Figure 1.

|  |  |  |
| --- | --- | --- |
| Close up of a plant  AI-generated content may be incorrect.   1. Bacteria | A close up of a leaf  AI-generated content may be incorrect.   1. Fungi | A hand holding a leaf  AI-generated content may be incorrect.   1. Nematode |
| A hand holding a leaf  AI-generated content may be incorrect.   1. Pest | A close-up of a leaf  AI-generated content may be incorrect.   1. Phytopthora | A close up of a plant  AI-generated content may be incorrect.   1. Virus |

## FIGURE 1. The given 6 disease classes

In terms of data processing, normalization, augmentation and segmentation has been implemented. The dataset had been manually segmented into its referring classes for further model implementations. The pixel values were transformed from the range of [0, 255] to [0,1]. This helps stabilize and speed up the training of the model which helps the model process more effectively. Lowering the pixel values to a fixed and smaller range allows the model to process the information more efficiently [17]. In an effort to make the model more robust and improve its generalization ability, data augmentation methods were employed on the training dataset. In particular, random augmentation was used to apply more transformations to the data. The manual setting for random rotation was specified to be 0.2 meaning images can be rotated randomly within certain angle geometries.

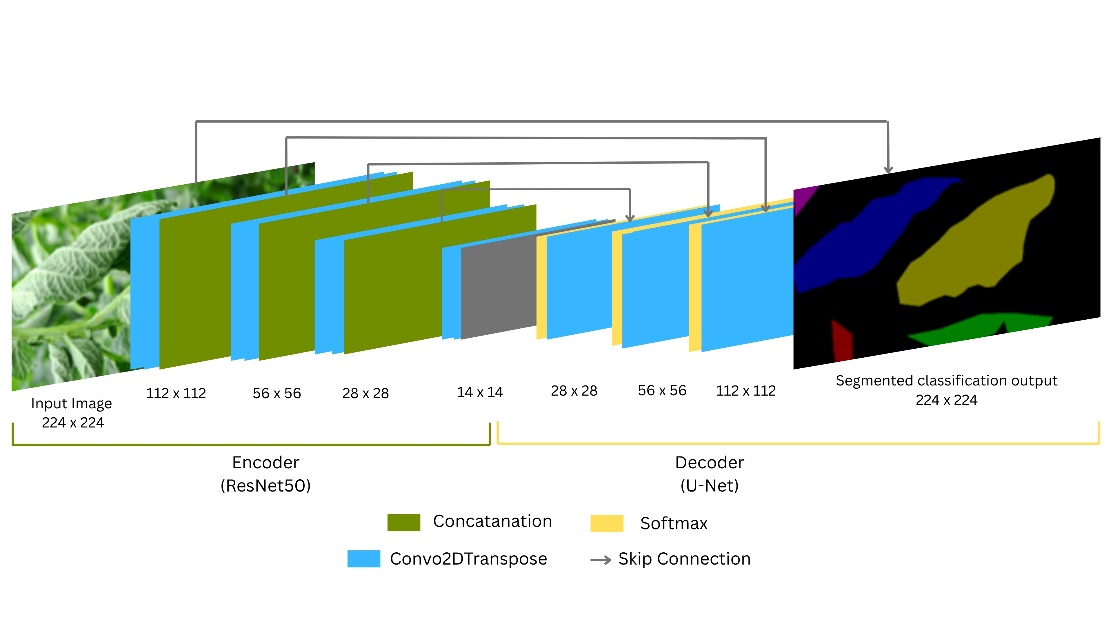
## Proposed Method

The LeafVision architecture uses a modified U-Net structure with image classification capabilities based on a pretrained ResNet50 encoder which extracts multi-level features from RGB images of dimensions 224×224. Spatial context is retained by skip connections from four essential layers of ResNet50 spatial context retention. The obtained bottleneck features are sent to a decoder that performs up-sampling through transposed convolutions and concatenation with corresponding encoder outputs. Additional feature refinement is accomplished through further convolution, followed by global average pooling and class prediction done by a dense SoftMax layer. The training phase used a learning rate of 0.0001 and applied early stopping at 100 epochs. This multi-purpose architecture leverages deep representational capabilities of ResNet50 for feature extraction and U-Net’s spatial reconstruction up-classification imaging data of complex nature. The model, shown in Figure 2, implements an image classification task using a U-Net inspired encoder-decoder architecture with ResNet50 as the encoder. The constraining output or bottleneck representation comes from conv5 while feature maps from ResNet are captured at four salient points; conv1, conv2, conv3 and conv4. The decoder is made of four transposed convolutional layers with up-sampling and skip connections from the encoder, these additively enhance the resolution at each up-sampling stage.

For classification, a Global Average Pooling layer followed by a few dense layers is utilized and the output is passed through a SoftMax layer to generate probabilities for the different classes. The model is trained using Adam optimizer with sparse categorical cross-entropy loss. The formula of sparse categorical cross-entropy loss is given:

(1)

For the ResNet50 encoder part, hyperparameters were tuned in pursuit of finding the best possible encoding architecture without overfitting to the data. The hyperparameters mentioned in Table 1, are the optimal ones found through rigorous testing.



**FIGURE 2**. Proposed LeafVision architecture

The proposed hyper-tuned system, as shown in Table 1, outperforms conventional hybrid models by employing a powerful encoder paired with a good decoder. The bridge layer from the deepest part of ResNet50 captures image features, providing context for the entire image and aids the overall imaging result. The output is refined by stepwise increasing the resolution in the decoder sections, so the more accurate details can be produced. The U-Net consists of global average pooling followed by a few dense layers for an efficient prediction. This Improved ResNet+U-Net provides more dependable results by balancing learning on deeper network layers and detailed image reconstruction.

**TABLE 1.** ResNet and U-Net hyperparameters

|  |  |  |  |
| --- | --- | --- | --- |
| **ResNet Encoder** | | **U-Net Decoder** | |
| **Component** | **Setting** | **Component** | **Setting** |
| Base Model | ResNet50 | Decoder Filters | 512 → 256 → 128 → 64 (Conv2DTranspose + Conv2D) |
| Weights | imagenet (pretrained) | Post-Upsample Pooling | GlobalAveragePooling2D |
| Include Top | False (no classification head) | Dense Layer | 128 units, ReLU |
| Input Tensor | Input (shape= (224, 224, 3)) | Output Layer | n\_classes units, Softmax |
| Skip Connection Layers | conv1\_relu (64 filters) | Optimizer | Adam |
| conv2\_block3\_out (256 filters) | Learning Rate | 1.00E-05 |
| conv3\_block4\_out (512 filters) | Loss Function | SparseCategoricalCrossentropy (from\_logits=False) |
| conv4\_block6\_out (1024 filters) | Metrics | Accuracy |
| Bridge Layer | conv5\_block3\_out (2048 filters) | Callback - Checkpoint | Save best model by val\_accuracy |

# Result

The dataset consists of 3076 images from six diseases: Bacteria, Fungi, Nematode, Pest, Phytophthora, and Virus; and the healthy class. Since the images are captured in uncontrolled environments, the differences in lighting conditions, background, orientation of the leaves, and scaling of the images significantly challenge classification. All evaluation metrics were considered; however, the most complex was U-Net model which outperformed the counterpart by a significant margin on state-of-the-art metrics. A detailed comparison with two baseline models, VGG16 and standalone ResNet50, is presented in Table 2.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TABLE 2.** Evaluation criteria of the models | | | | |
| **Model** | **Accuracy** | **F1 Score** | **Recall** | **Precision** |
| VGG | 0.90 | 0.912 | 0.908 | 0.90 |
| ResNet50 | 0.927 | 0.926 | 0.927 | 0.93 |
| **LeafVision** | **0.979** | **0.979** | **0.9789** | **0.98** |

The evaluation for the LeafVision model resulted in undermentioned outcomes precision: 0.98, F1 score: 0.979, accuracy: 0.979 and recall: 0.9789. Hence, the U-Net model achieved the highest performance and surpassed the VGG16 and ResNet50 baselines significantly.

Figure 3 displays the confusion matrix, which illustrates the model's excellent classification accuracy for all six disease and healthy classes with only minimal misclassifications occurring for each class. This evidences the model’s accuracy and attention to subtle visual differences across categories of diseases. Figure 4 displays the training and validation accuracy as well as loss curves for 100 epochs. The training and validation accuracy graphs both increases, although validation accuracy has some fluctuations. Also, the validation loss shows spikes, but the overall trend is decreasing. Such behavior is normal in datasets gathered from uncontrolled environmental conditions.

From Figure 4 and Figure 5 it is apparent that the proposed model is not over-fitted for the dataset and that accurate classification of diseases with high confidence score is obtained despite the irregularity of the training and validation accuracy, which emphasizes the generalization capability of the model.

A screenshot of a graph

AI-generated content may be incorrect.

**FIGURE 3**. Confusion matrix of LeafVision

A graph of a training and validation accuracy

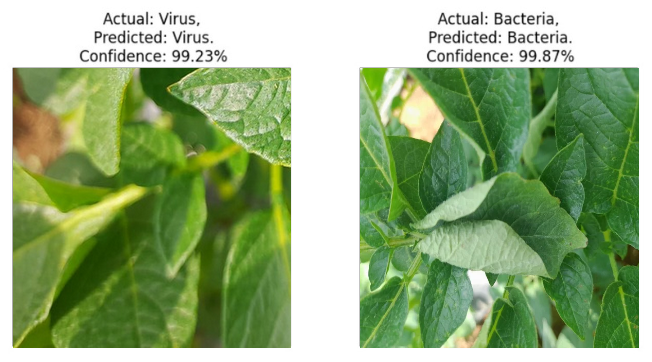
AI-generated content may be incorrect.

a) Training and validation accuracy b) Training and validation loss

**FIGURE 4**. Training and validation a) Accuracy and b) Loss curve

# Discussion

This study proposed LeafVision, a ResNet+U-Net DL model and ratified the effectiveness of the proposed U-Net model, utilizing a ResNet50 encoder for potato leaf disease classification, especially on images taken under uncontrolled conditions. Unlike most studies that use datasets curated under very controlled conditions, the dataset in this research spans a wide range of variables including varying lighting, background, and scale. It was noted that the U-Net model with the capability of feature extraction by ResNet50 and preservation of spatial detail through skip connections for lower layers, solves these issues. This undoubtedly contributed toward the performance improvement of the model, which achieved 97.9% accuracy compared to the 90% and 92.7% accuracies obtained by VGG16 and ResNet50 as noted in Table 2. Our proposed model achieved an **F1-score of 97.9%,** a **recall of 97.89%,** and a **precision of 98% which highlighted the model has a strong and balanced classification performance. The high precision and recall value indicate the model’s reliability and robustness in real world scenario specially** in uncontrolled environment where image quality and conditions are varied. The model surpasses expectations for its ability to classify seven distinct categories including 6 groups of disease which is considered a higher range of classes in comparison to two to four found in similar reviewed literature. These categories- Bacteria, Fungi, Nematode, Pest, Phytophthora and Virus, are quite difficult to discriminate due to the underlying subtle differences in leaf symptoms. Nonetheless, the proposed model was able to differentiate them with high precision. Noticeable classification errors are not apparent in any of the matrices, which implies balanced accuracy in all categories. Also, the training and validation curves showcase the learning process to be steady and uniform. The training and validation accuracy measures as well as loss values improve without incessantly increasing, indicates is a sign the model has indeed generalized well to the problem in hand, which is good especially in the context of the variability in the dataset. From these results it is clear, LeafVision is effective in potato leaf disease classification under practical conditions. Integrated multi-scale feature extraction and detail spatial retention with generalization in the multi-scale add promising approach for usage in precision agriculture where acute reliable detection of diseases can assist in effective management of crops.



1. (b)

**FIGURE 5**. Visualization of model predictions on validation images showing (a) actual and (b) predicted classes with confidence

# Conclusion

This paper presented LeafVision, a hybrid model which have the power of U-Net architecture with the feature extraction capabilities of a ResNet50 backbone design for multi-class classification of potato leaf diseases using images captured in uncontrolled field settings. The model addresses a critical challenge in agriculture technology, detecting plant diseases accurately from real-world images characterized by variable lighting, complex backgrounds, different leaf orientations, and diverse image qualities. The proposed model achieves high accuracy to detect six classes of disease where multiple diseases are visually similar in many cases with overlapping symptoms such as discoloration, lesionns, and wilting which makes precise classification particularly difficult. Despite this, LeafVision was able to accurately differentiate between these classes, showcasing its ability to capture fine-grained and disease-specific features. The combination of ResNet+**U-Net encoder-decoder structure with skip connections** and **deep residual learning** plays an important role in the model’s high accuracy. The adaptation of specialization segmentation architecture in classification processed indeed preserved some details which other standard and hybrid CNN approaches do not capture an important factor for accuracy in disease detection. Because the data was not altered much, this strong performance enables the model to be easily used in portable or drone-based diagnostic systems for real-time precision agriculture monitoring. Further development will be aimed at model size reduction for integrated devices, broader targets, and user-friendly systems for enhanced precision agriculture and other crops and diseases.

# Reference

1. V.L. Adluri, P.S. Reddy, S. Palthya, S. Jampani, and S.P. Gandhari, “Potato Leaf Disease Detection and Classification Using VGG16,” in *Proceedings of International Joint Conference on Advances in Computational Intelligence*, edited by M.S. Uddin and J.C. Bansal, (Springer Nature, Singapore, 2024), pp. 147–158.
2. A.N. Jui, N. Dey, K.N.-E.-A. Siddiquee, and S.M. Shawon, “Botanika: An IoT-Enabled Solar Powered Soil Profiling Sensor Probe for Rooftop Gardening,” in *2024 3rd International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, (2024), pp. 1–6.
3. S.M. Shawon, A.N. Adnan, K.M.S. Noor, Z. Kabir, and A. Chakraborty, “Design and Development of Solar Powered IoT Enabled Multi-layer Soil Moisture Probe,” in *2024 27th International Conference on Computer and Information Technology (ICCIT)*, (2024), pp. 3456–3461.
4. U.Y. Tambe, A. Shobanadevi, A. Shanthini, and H.-C. Hsu, “Potato Leaf Disease Classification using Deep Learning: A Convolutional Neural Network Approach,” (2023).
5. M. Thalor, M. Pathak, R. Jamadar, and S. Khan, “Potato Leaf Insight: An AI-Driven System for Detection of Potato Leaf Disease,” Curr Agri Res Jour **12**(3), 1135–1141 (2025).
6. J. Rashid, I. Khan, G. Ali, S.H. Almotiri, M.A. AlGhamdi, and K. Masood, “Multi-Level Deep Learning Model for Potato Leaf Disease Recognition,” Electronics **10**(17), 2064 (2021).
7. S.M. Shawon, F. Barua Ema, A.K. Mahi, and Md. Mohsin Sarker Raihan, “Crop Yield Prediction: Robust Machine Learning Approaches for Precision Agriculture,” in *2023 26th International Conference on Computer and Information Technology (ICCIT)*, (IEEE, Cox’s Bazar, Bangladesh, 2023), pp. 1–6.
8. S.M. Shawon, N.I. Neha, A.N. Jui, N. Dey, and Md.M.S. Raihan, “AgriAI: Machine Learning Frameworks for Tailored Crop Recommendations using Soil Nutrient Parameters,” in *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS)*, (IEEE, Dhaka, Bangladesh, 2024), pp. 1–6.
9. S.M. Shawon, F.B. Ema, A.K. Mahi, F.L. Niha, and H.T. Zubair, “Crop yield prediction using machine learning: An extensive and systematic literature review,” Smart Agricultural Technology **10**, 100718 (2025).
10. D. Tiwari, M. Ashish, N. Gangwar, A. Sharma, S. Patel, and S. Bhardwaj, “Potato Leaf Diseases Detection Using Deep Learning,” in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, (IEEE, Madurai, India, 2020), pp. 461–466.
11. Md.A. Iqbal, and K.H. Talukder, “Detection of Potato Disease Using Image Segmentation and Machine Learning,” in *2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, (IEEE, Chennai, India, 2020), pp. 43–47.
12. C.C. Bonik, F. Akter, Md.H. Rashid, and A. Sattar, “A Convolutional Neural Network Based Potato Leaf Diseases Detection Using Sequential Model,” in *2023 International Conference for Advancement in Technology (ICONAT)*, (IEEE, Goa, India, 2023), pp. 1–6.
13. J. Chen, X. Deng, Y. Wen, W. Chen, A. Zeb, and D. Zhang, “Weakly-supervised learning method for the recognition of potato leaf diseases,” Artif Intell Rev **56**(8), 7985–8002 (2023).
14. Z. Saeed, M. Urooj Khan, A. Raza, N. Sajjad, S. Naz, and A. Salal, “Identification of Leaf Diseases in Potato Crop Using Deep Convolutional Neural Networks (DCNNs),” in *2021 16th International Conference on Emerging Technologies (ICET)*, (IEEE, Islamabad, Pakistan, 2021), pp. 1–6.
15. M. Radwan, A.A. Alhussan, A. Ibrahim, and S.M. Tawfeek, “Potato Leaf Disease Classification Using Optimized Machine Learning Models and Feature Selection Techniques,” Potato Res. **68**(2), 897–921 (2025).
16. N.H. Shabrina, S. Indarti, R. Maharani, D.A. Kristiyanti, Irmawati, N. Prastomo, and T. Adilah M, “A novel dataset of potato leaf disease in uncontrolled environment,” Data in Brief **52**, 109955 (2024).
17. M.Y. Xin, L.W. Ang, and S. Palaniappan, “A Data Augmented Method for Plant Disease Leaf Image Recognition based on Enhanced GAN Model Network,” *Journal of Informatics and Web Engineering* **2**(1), 1–12 (2023).