

ResNet 50 and Inception V3-based Deep Learning Classification of Potato Leaf Diseases

Anurag Shrivastava¹

¹Assistant Professor (Instrumentation), BRSM College of Agricultural Engineering and Technology, Mungeli

ABSTRACT

Potato is one of the most widely cultivated and consumed crops in India and has substantial demand in the global market. However, various diseases significantly affect potato quality and yield, leading to considerable economic losses. Diseases such as early blight, late blight, common scab, and potato leafroll virus adversely impact both the quality and quantity of potato production. Early and accurate disease detection is therefore essential for effective crop management. Computer vision-based automated systems can assist farmers in identifying diseases at an early stage, enabling timely intervention. In this study, a deep learning-based approach is proposed for the detection and classification of potato leaf diseases. A total of 2,150 leaf images were collected from an online database and categorized into three classes: early blight, late blight and healthy leaves. A ResNet 50 and Inception V3 deep learning models were employed for disease classification, achieving an accuracy of 94.33% for the Inception V3 deep learning architecture. The results demonstrate the effectiveness of deep learning techniques for potato disease detection and highlight their potential to reduce crop losses and improve agricultural productivity.

Keywords: *Potato leaf disease, deep learning, classification, ResNet 50, Inception V3.*

1. INTRODUCTION

Agriculture is a fundamental pillar of the Indian economy and plays a crucial role in ensuring food security, employment generation and rural development. According to recent agricultural statistics, a significant proportion of India's population depends directly or indirectly on agriculture and allied activities for their livelihood [1]. Among the various horticultural and vegetable crops cultivated in India, potato is one of the most important food crops due to its high nutritional value, adaptability to diverse climatic conditions and widespread consumer acceptance. Potato ranks among the world's leading food crops after rice, wheat, and maize and contributes significantly to global food production [2].

Despite their importance, potato crops are highly vulnerable to various biotic and abiotic stresses that adversely affect growth, quality and yield. Diseases caused by fungi, bacteria, viruses and nematodes can substantially reduce crop productivity and market value.

Common potato diseases include early blight, late blight, common scab, blackleg disease, bacterial etc. Among all potato diseases, early blight and late blight are considered particularly destructive. Early blight typically manifests as dark brown lesions with concentric rings on older leaves, leading to premature defoliation and reduced photosynthetic activity, whereas late blight, historically known for causing the Irish Potato Famine, is one of the most devastating diseases affecting potato crops. It spreads rapidly under favorable environmental conditions and can destroy entire fields within a short period if not detected and controlled promptly [3,4]. Traditionally, disease identification relies on manual visual inspection performed by farmers, agricultural officers or plant pathologists. Although expert-based diagnosis can be effective, it suffers from several limitations. First, the process is labor-intensive and time-consuming, especially when monitoring large agricultural fields. Second, accurate diagnosis requires extensive expertise and experience, which may not always be available in rural or resource-constrained regions. Third, visual symptoms of different diseases often appear similar during the early stages of infection, increasing the likelihood of misclassification and delayed treatment [5]. Such delays may allow diseases to spread rapidly, resulting in greater crop damage and economic losses. Therefore, the development of automated and intelligent disease detection systems has become an active area of research in agricultural engineering and computer science.

Recent advances in artificial intelligence (AI), machine learning (ML), deep learning (DL) and computer vision technologies have revolutionized agricultural monitoring and decision-making processes. Several image-based machine learning approaches have been successfully applied to plant disease detection. Traditional classifiers rely heavily on manually extracted features. Although effective, their performance may decline when dealing with complex image datasets characterized by variations in illumination, background, scale and disease severity [6].

To overcome these limitations, deep learning techniques have gained significant attention in recent years. Deep learning models automatically learn hierarchical feature representations directly from raw image data, eliminating the need for manual feature engineering. Among various deep learning architectures, Convolutional Neural Networks (CNNs) have achieved remarkable success in image recognition and classification tasks. Popular CNN

architectures such as ResNet and Inception deep learning models have been widely employed for plant disease identification with impressive results [7,8].

The availability of publicly accessible datasets has further accelerated research in automated plant disease detection. Numerous studies have reported high classification accuracies using deep learning approaches for identifying potato diseases [9]. Specifically, potato leaf disease detection using CNN-based models has demonstrated high reliability and robustness under controlled imaging conditions. Early disease detection not only reduces crop losses but also minimizes excessive pesticide usage, lowers production costs, and promotes sustainable agricultural practices [7,8].

Motivated by these advancements, the present study investigates the application of deep learning techniques for the classification of potato leaf diseases. The proposed framework aims to distinguish between healthy leaves and diseased leaves affected by major potato diseases (early and late blight) using image-based analysis. By leveraging modern computer vision and artificial intelligence techniques, the study seeks to develop an efficient, accurate, and scalable disease detection system that can assist farmers and agricultural stakeholders in improving crop health management. The successful implementation of such systems has the potential to enhance potato productivity, reduce economic losses, support precision agriculture initiatives and contribute to sustainable agricultural development.

2. MATERIALS AND METHODS

Figure 1 shows a block diagram of the proposed methodology, it includes the following four main steps: data acquisition, data pre-processing, data augmentation, and classification using ResNet 50 and Inception V3 deep learning models. Approximately 2150 images of potatoes were employed from Kaggle, specifically the Plant Village Dataset. This dataset includes images categorized into classes such as Early blight, Late blight, and Healthy. Some of the images from the dataset are displayed in Figures 2 and 3. Data preprocessing seeks to enhance data quality to achieve precise results from the training model. Initially, the goal is to reduce image noise, and images with too much noise are discarded.

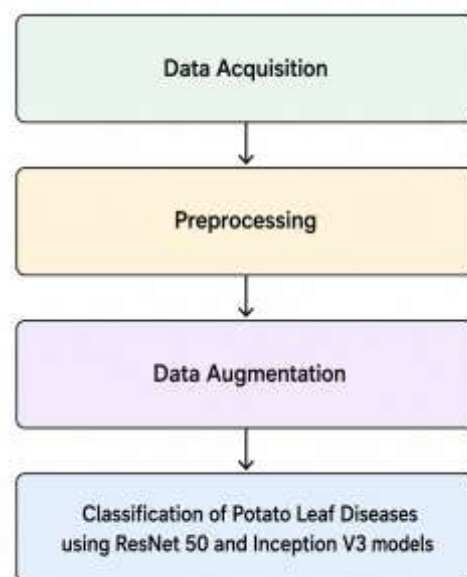


Figure 1: Block diagram of study



Figure 2. Early Blight: Symptoms in this disease are leaves with brown, non-glossy, dead spots, concentric rings in a target-board pattern.



Figure 3. Healthy leaves

Data preprocessing seeks to enhance data quality to achieve precise results from the training model. Initially, the goal is to reduce image noise, and images with too much noise are discarded. The images obtained come in different dimensions, so they are adjusted to 256x256 pixels to ensure uniformity in dataset inputs. Deep Learning demands significantly more data than shallow networks in machine learning. A common issue in machine learning is insufficient training data and class-imbalanced data distributions. Data augmentation is a method of altering data while preserving its core characteristics. Various techniques are employed in data augmentation, ranging from basic image transformations such as flipping, rotating, enlarging, and cropping to

histogram-based methods. In this research, only basic transformation techniques such as rotating and cropping were utilized. Finally, the dataset were applied to deep learning models (ResNet 50 and Inception V3) to classify among health and diseased leaves (early and late blight). Deep learning model architectures: The ResNet-50 model is a powerful Convolutional Neural Network (CNN) with 50 layers. Developed by He et al., it tackles the issue of vanishing gradients in deep networks by employing residual learning and identity shortcut connections. These "skip connections" enable the network to circumvent certain layers, maintaining a stable gradient flow and facilitating effective feature extraction in deeper layers without compromising optimization. The architecture begins with a 7×7 convolutional layer, followed by a sequence of residual blocks. Each block incorporates bottleneck designs using 1×1 , 3×3 , and 1×1 convolutions to minimize computational demands.[7]

Inception-v3 architecture represents a highly refined Convolutional Neural Network (CNN). This version enhances earlier models by adopting an efficient design that emphasizes factorized convolutions and strong regularization techniques. Departing from conventional sequential structures, Inception-v3 employs unique Inception modules that execute parallel convolutional operations with filters of different dimensions (such as 1×1 , 3×3 , and 5×5) alongside max pooling within a single layer. To minimize computational demands and avoid overfitting, larger convolutions are broken down into smaller, sequential ones (for instance, a 5×5 convolution is split into two 3×3 convolutions, and further into asymmetric 1×3 and 3×1 convolutions). The architecture also integrates RMSProp optimization, label smoothing, and auxiliary classifiers to stabilize the training process and preserve spatial data, culminating in a global average pooling layer followed by a fully connected softmax layer for generating final predictions.[8]

ResNet-50 utilizes residual learning blocks with identity shortcut connections to mitigate the vanishing gradient problem in deep neural networks. In contrast, Inception-v3 employs a multi-scale processing approach by integrating factorized convolutions and asymmetric kernels to capture spatial hierarchies while maintaining computational efficiency. These architectural distinctions reflect a fundamental trade-off between the depth-driven feature representation of ResNet-50 and the width-optimized, modular topology characteristic of the Inception-v3 framework. Empirical evaluations demonstrate that while ResNet-50 often excels in tasks requiring high-resolution semantic consistency, Inception-v3 provides superior performance in scenarios involving high intra-class variance through its diversified filter bank. Furthermore, the computational overhead of ResNet-50 scales primarily with network depth, whereas Inception-v3 incurs higher parameter complexity due to its intricate inception modules. Consequently, the selection between these architectures necessitates a rigorous assessment of the specific hardware constraints

and the underlying topological complexity inherent in the target dataset. Moreover, fine-tuning these models requires careful consideration of the receptive field requirements, as ResNet-50's bottleneck design is inherently better suited for standard spatial hierarchies, while Inception-v3's modularity facilitates more nuanced feature extraction for multi-scale object detection tasks. This divergence in design philosophy suggests that ResNet-50 serves as a robust baseline for feature extraction in classification-heavy architectures, whereas Inception-v3 provides the structural flexibility necessary for complex feature fusion in vision-based systems. Ultimately, the integration of these models into production pipelines remains highly dependent on the trade-off between the depth-wise stability offered by residual connections and the multi-granular feature capture enabled by the inception architecture.

Hyperparameter Tuning

Table 2 shows the hyperparameter tuning for training ResNet 50 and Inception V3 deep learning models. These settings control the learning process, optimization, validation, and hardware utilization during model training.

Table 1: Hyperparameter initialization for ResNet 50 and Inception V3 architectures.

Parameter	Value	Description
Solver	Adam	An optimization algorithm used to update network weights. Adam combines the advantages of Momentum and RMS Prop for faster convergence.
Initial Learning Rate	0.0001	Determines the step size during weight updates. A small value helps achieve stable learning.
Mini Batch Size	16	Number of images processed before updating network weights.
Maximum Epochs	11	Total number of complete passes through the training dataset.
Validation Frequency	50	Model performance is evaluated on validation data after every 50 iterations.

Performance evaluation: To rigorously assess the performance of our deep learning models, we employed a comprehensive suite of standard evaluation metrics derived from the confusion matrix. The overall accuracy was calculated to measure the proportion of correctly predicted instances out of the total dataset. However, to account for potential class imbalances and ensure a nuanced evaluation, we also analyzed precision, recall,

and the F1-score. Precision quantifies the model's exactness by measuring the ratio of true positive predictions to all positive predictions, while recall (sensitivity) evaluates the model's completeness by capturing the ratio of true positives to all actual positive instances. The F1-score was computed as the harmonic mean of precision and recall, providing a single, balanced metric of overall classification performance. The expressions for the evaluation matrix parameters are given by equations 1 to 5 [10].

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

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

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{Specificity} = \frac{TN}{FP+TN} \quad (4)$$

$$\text{F1 Score} = \frac{2 \times [\text{Precision} \times \text{Recall}]}{[\text{Precision} + \text{Recall}]} \quad (5)$$

3 RESULTS

The models were tested based on 80-20 percent split cross-validation. This means 80 percent of the data from the dataset was used for training and rest 20 % used for validation. The experimental results show the performance of ResNet 50 and Inception V3 deep learning architectures for potato leaf disease classification. Models were tested for variable batch size and the results are reported in Tables 2 and 3. The training, validation and loss function plot for batch size of 64 for both the models are given by Figures 4 and 5.



Figure 4: Training and validation plot for ResNet 50 model with batch size 64.

Both ResNet50 and Inception V3 achieved high accuracy, but they responded to batch size changes differently. The Inception V3 performed best at the largest batch size. With a batch size of 64, it achieved its highest validation accuracy of 94.33% and its lowest validation loss of 0.0418. This suggests that larger batch sizes provide more stable gradient steps for this architecture, helping it learn cleaner patterns from potato leaf images. ResNet50 peaked earlier. It

reached its best validation accuracy of 91.50% at a batch size of 32. When the batch size increased to 64, its training accuracy continued to rise to 96.43%, but its validation accuracy dropped slightly to 91.33%. This minor gap indicates a small amount of overfitting, where ResNet50 began memorizing the training details rather than generalizing to new leaf samples.

Table 2 - Performance of ResNet50 for different batch sizes.

Batch Size	Learning Rate	Training Accuracy (%)	Validation Accuracy (%)	Validation Loss
16	0.0001	89.50	87.20	0.135
32	0.0001	93.80	91.50	0.118
64	0.0001	96.43	91.33	0.082

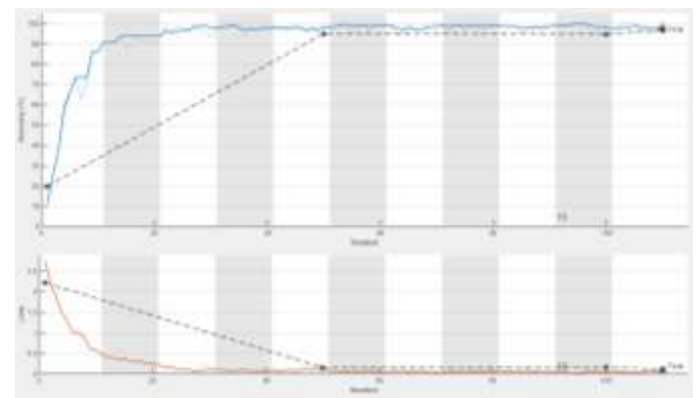


Figure 5: Training and validation plot for Inception V3 model with batch size 64.

Inception V3 consistently beat ResNet50 in peak performance and error reduction. At a batch size of 64, the Inception V3 validation loss (0.0418) was nearly half that of ResNet50 (0.0820). This lower loss means the Proposed Model makes predictions with much higher confidence. ResNet50 remains a strong, reliable baseline, but the Proposed Model proves better suited for capturing the fine visual details, such as early blight spots or late blight lesions, needed for accurate potato leaf disease classification.

Table 3 - Performance of Inception V3 for different batch sizes.

Batch Size	Learning Rate	Training Accuracy (%)	Validation Accuracy (%)	Validation Loss
16	0.0001	91.43	85.67	0.1754
32	0.0001	96.50	92.84	0.1212
64	0.0001	97.92	94.33	0.0418

This superior performance is further supported by Inception V3's exceptional **sensitivity (98.57%)** and **specificity (97.95%)**, which indicate its strong ability to minimize both missed positive targets and false alarms. In contrast, ResNet 50 achieved slightly lower values, with a sensitivity of 94.81% and a specificity of 94.62%. Furthermore, Inception V3 demonstrated superior overall prediction balance, securing a **precision of 97.33%** and an **F1 Score of 97.26%**, whereas ResNet 50 yielded a precision of 96.17% and an F1 Score of 95.49% as shown in Table 4.

Table 4: Evaluation Matrices for ResNet 50 Deep Learning Architecture

Deep Learning Architecture	ResNet 50	Inception V3
Accuracy	96.43	97.92
Precision	96.17	97.33
Sensitivity	94.81	98.57
Specificity	94.62	97.95
F1 Score	95.49	97.26

Conclusion

This study evaluated deep learning models to classify potato leaf diseases using different training setups. The results prove that tuning the batch size is vital for maximizing classification success.

The Proposed Model with a batch size of 64 and a learning rate of 0.0001 delivered the best results, yielding a 94.33% validation accuracy and a 0.0418 validation loss. This outshone the best ResNet50 configuration, which reached 91.50% accuracy.

These findings show that our optimized deep learning approach can reliably identify potato plant diseases. In the future, this model could be built into mobile apps or drone systems to help farmers spot crop diseases early in the field, reducing pesticide use and protecting crop yields.

4. REFERENCES

- [1] Singh, S. (2026). Agriculture in the Indian economy: growth, employment, and structural challenges. *Social Science Journal for Advanced Research*, 6(1), 73–77. <https://doi.org/10.54741/ssjar/6.1.2026.321>.
- [2] Dolničar, P. (2021). Importance of potato as a crop and practical approaches to potato breeding. *Methods in Molecular Biology*, 2354, 3–20. https://doi.org/10.1007/978-1-0716-1609-3_1.
- [3] Pooja, Chauhan, P., Kumar, A., Kumar, A., Rithesh, L., Kumar, A., & Kumar, A. (2025). Recent insights into potato dry rot an emerging Disease: Focusing on pathogen diversity, host-pathogen interactions, and management strategies. *Microbial Pathogenesis*, 207, 107866. <https://doi.org/10.1016/j.micpath.2025.107866>.
- [4] Farooq, M. S., Kamran, A., Raza, S. A., Wasiq, M. F., Hassan, B., & Herzog, N. J. (2025). Advancing early blight detection in potato leaves through ZeroShot learning. *Journal of Imaging*, 11(8), 256. <https://doi.org/10.3390/jimaging11080256>.
- [5] Sharma, A., Singh, V., Sharma, C., Ansari, G., Kumar, S., & Joshi, K. (2024). *Potato Leaf Disease Detection and Classification Using Deep Learning Technique* (pp.505-509) <https://doi.org/10.1109/ictacs62700.2024.10840790>.
- [6] Sinamenye, J. H., Chatterjee, A., & Shrestha, R. (2025). Potato plant disease detection: leveraging hybrid deep learning models. *BMC Plant Biology*, 25(1), 647. <https://doi.org/10.1186/s12870-025-06679-4>.
- [7] Shaheed, K., Qureshi, I., Abbas, F., Jabbar, S., Abbas, Q., Ahmad, H., & Sajid, M. Z. (2023). EfficientRMT-NeT—An efficient RESNET-50 and Vision Transformers approach for classifying potato plant leaf diseases. *Sensors*, 23(23), 9516. <https://doi.org/10.3390/s23239516>.
- [8] Nur, K. N. A., Addyna, N. I., Windarto, A. P., Wanto, A., & Poningsih, P. (2025). OPTIMIZATION OF THE INCEPTIONV3 ARCHITECTURE FOR POTATO LEAF DISEASE CLASSIFICATION. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 10(4), 849–858. <https://doi.org/10.33480/jitk.v10i4.6554>.
- [9] A. Suryavanshi, S. Tanwar, V. Kukreja, A. Choudhary and S. Chamoli, "An Integrated Approach to Potato Leaf Disease Detection Using Convolutional Neural Networks and Random Forest," *2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems*

(ICSES), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICSES60034.2023.10465557.

[10] Asif, M., Abbas, T., Iqbal, S. *et al.* MapReduce-based deep learning framework for potato leaf disease detection in sustainable precision agriculture. *Sci Rep* **16**, 1607 (2026). <https://doi.org/10.1038/s41598-025-30940-3>.