Automated Potato Leaf Disease Identification Using Deep Learning and Image Processing

Sangeeta Jana Mukhopadhyay
Department of Electronics and Communication Engineering, Brainware University, Barasat, Kolkata, West Bengal, India

Abstract: One of the most extensively grown crops in the world, potatoes are essential to food security. However, potato production is severely affected by several leaf diseases, including Late Blight, Early Blight, Mosaic Virus (PVY), and Black Leg. Conventional diagnostic techniques rely on manual inspection, which is labour-intensive, prone to error, and unsuitable for extensive farming regions. In this paper, a Convolution Neural Network (CNN)-based automated potato leaf disease detection system based on deep learning and image processing is presented. A labeled dataset of five classes (Late Blight, Early Blight, Mosaic Virus, Black Leg, and Healthy) was collected and pre-processed using advanced augmentation and contrast-enhancement techniques. The proposed hybrid CNN achieved an overall accuracy of 95.8%, outperforming SVM, Random Forest, and the baseline CNN. ROC curves, confusion matrix analysis, and performance metrics confirm the model's robustness. A lightweight, user-friendly GUI was developed to provide real-time disease prediction and recommendations for field applications. The system enables early detection, reduces misdiagnosis, and supports sustainable potato cultivation.

Keywords: Potato diseases, Late Blight, Early Blight, CNN, Image Processing, Deep Learning, Precision Agriculture

1. Introduction

One of the most widely grown root crops in the world, potatoes (Solanum tuberosum) are essential to global food security. Its high adaptability, nutritional value, and demand across food processing industries make it essential for both subsistence and commercial farming. However, potato production is severely threatened by several foliar diseases, including Late Blight, Early Blight, Potato Virus Y (PVY), and Black Leg, which collectively cause significant yield and economic losses. In many major potato-growing regions, these diseases can cause 30–80% yield reduction when early detection and management are not implemented [1], [2].

Traditional plant disease identification relies on manual inspection, which is labour-intensive, inconsistent, and prone to human error due to subjectivity and varying levels of expertise. In the early phases of disease development, when visual symptoms are modest or resemble environmental stress responses, such methods are frequently ineffective for large-scale surveillance and provide limited accuracy [3]. These drawbacks emphasize the necessity of automated, intelligent, and real-time disease detection systems, particularly those that can function in field settings with complicated backgrounds and fluctuating lighting [4].

The diagnosis of agricultural diseases has been dramatically enhanced by recent developments in Artificial Intelligence (AI), especially Deep Learning (DL) and Computer Vision. Because they can automatically learn spatial and textural features, Convolutional Neural Networks (CNNs) have become the most potent architecture for image-based illness diagnosis [5], [6]. By eliminating manual feature extraction and achieving greater generalisation accuracy, CNN-based techniques outperform conventional machine learning methods [7].

Several researchers have successfully implemented deep learning models for potato disease classification in both controlled and field conditions. Transfer learning architectures such as ResNet50, VGG16, and Inception variants have achieved accuracy above 95% [8]–[10]. Improved performance has also been demonstrated using robust preprocessing, augmentation, and illumination normalization techniques [11], [12]. More advanced studies introduced texture-enhanced and hybrid CNN architectures to distinguish between visually overlapping symptoms, such as those seen in viral and fungal infections [13]–[15].

Although deep learning-based disease detection has achieved high recognition accuracy, challenges remain, including limited dataset variability, computational constraints on mobile devices, and symptom similarity during early stages [16], [17]. These factors motivate the development of lightweight, field-deployable CNN models capable of real-time multi-disease classification.

In order to overcome these obstacles, the current study suggests a user-interactive interface for real-time potato leaf disease diagnostics, an efficient pre-processing pipeline, and anoptimized hybrid CNN architecture. To promote sustainable potato production, the system aims to improve classification accuracy, reduce misdiagnosis, and help farmers make early decisions.

2. Diseases of Potato Leaves and Their Signs

Potato crops are particularly vulnerable to a range of foliar diseases that affect photosynthetic efficiency, canopy development, physiological function, and, ultimately, tuber yield. Since these symptoms directly affect the effectiveness of image-based deep learning models for feature extraction, it is crucial to understand the symptomatic patterns of each disease to design an automated classification system [18]. The four main potato illnesses taken into consideration in this study—Late Blight, Early Blight, Potato Virus Y (PVY/Mosaic Virus), and Black Leg—as well as the traits of healthy leaves, are described in more detail in this section. Figure 1 shows representative samples for each class.

2.1 Phytophthora infestansor late blight

The most destructive potato disease in the world, late blight has historically caused catastrophic famines. The oomycete Phytophthora infestans, which prefers cool, damp conditions, is the source of the illness. Small, pale-green water-soaked patches, usually at the tips or margins of leaves, are the first signs. These lesions rapidly enlarge into irregular, dark brown necrotic patches with a characteristic "burnt" appearance. Under high humidity, a fine, white mycelial growth can be observed on the undersides of infected leaves, representing the pathogen's sporulation structures. As disease severity increases, affected leaves collapse, shrivel, and detach prematurely, leading to complete canopy destruction. Accurate early diagnosis is essential to preventing epidemics and significant production losses due to the pathogen's rapid infection cycle and spore spread [19], [20], [21].

2.2 Early Blight (Alternaria solani)

Early Blight is a common fungal disease caused by Alternaria solani, typically appearing first on older leaves due to physiological stress or reduced nutrient mobility. The disease begins with small, round, brown spots that gradually expand and develop concentric rings, forming the classical “targetspot” pattern.

This ring-like structure is a key visual feature distinguishing Early Blight from other foliar diseases.

Lesions are often surrounded by chlorotic yellow halos, which result from the plant’s hypersensitive response. As disease progression continues, spots coalesce, leading to significant foliage deterioration and reduced photosynthetic capacity. Although Early Blight progresses more slowly than Late Blight, it remains economically damaging, especially under conditions of high temperature, leaf wetness, and nitrogen deficiency [22], [23], [24].

2.3 Potato Virus Y (PVY / Mosaic Virus)

Potato Virus Y (PVY) is among the most widespread viral diseases affecting solanaceous crops. Unlike fungal infections that produce necrotic lesions, PVY primarily alters pigment distribution. Its symptoms include mosaic mottling, alternating patches of dark and light green, leaf wrinkling, vein clearing, and upward leaf curling. Symptom severity varies with potato cultivar, virus strain, and environmental conditions. Mild infections may only show faint mosaic patterns, while severe infections lead to pronounced chlorosis, distorted leaves, stunted growth, and reduced tuber quality. The non-necrotic nature of PVY makes it visually distinct yet challenging for automated systems, as mosaic patterns can be subtle and resemble nutrient stress symptoms under field conditions [25], [26], [27].

2.4 Black Leg Disease

Black Leg is a bacterial disease primarily caused by Pectobacterium atrosepticum. While its most prominent symptoms appear on stems, leaves often reveal early visual indicators of infection. Initial symptoms include general yellowing (chlorosis), downward leaf curling, and loss of turgor.

As the disease progresses, the edges and tips of leaves may show dark brown to black discoloration.

The bacteria move through the plant's vascular tissues, interfering with the movement of nutrients and water, which eventually causes the plant to collapse, wilt its leaves, and turn black at the stem. In severe conditions, infected plants exhibit complete canopy degeneration. Because leaf symptoms can appear before stem discolouration becomes visible, early detection on foliage is crucial for timely disease management and prevention of tuber rotting in storage [28], [29].

2.5 Healthy Leaves

Healthy potato leaves present a uniform green colouration with smooth, undamaged lamina. Vein patterns are symmetrical, and the leaf surface retains a consistent texture without signs of necrosis, chlorosis, mosaic patterns, or deformation. Healthy leaves serve as the baseline class for the CNN model, enabling the system to distinguish between natural plant conditions and pathological deviations caused by biotic stress. Incorporating healthy samples ensures the model does not misclassify normal environmental variations, such as slight discoloration due to shading, as disease symptoms [30].

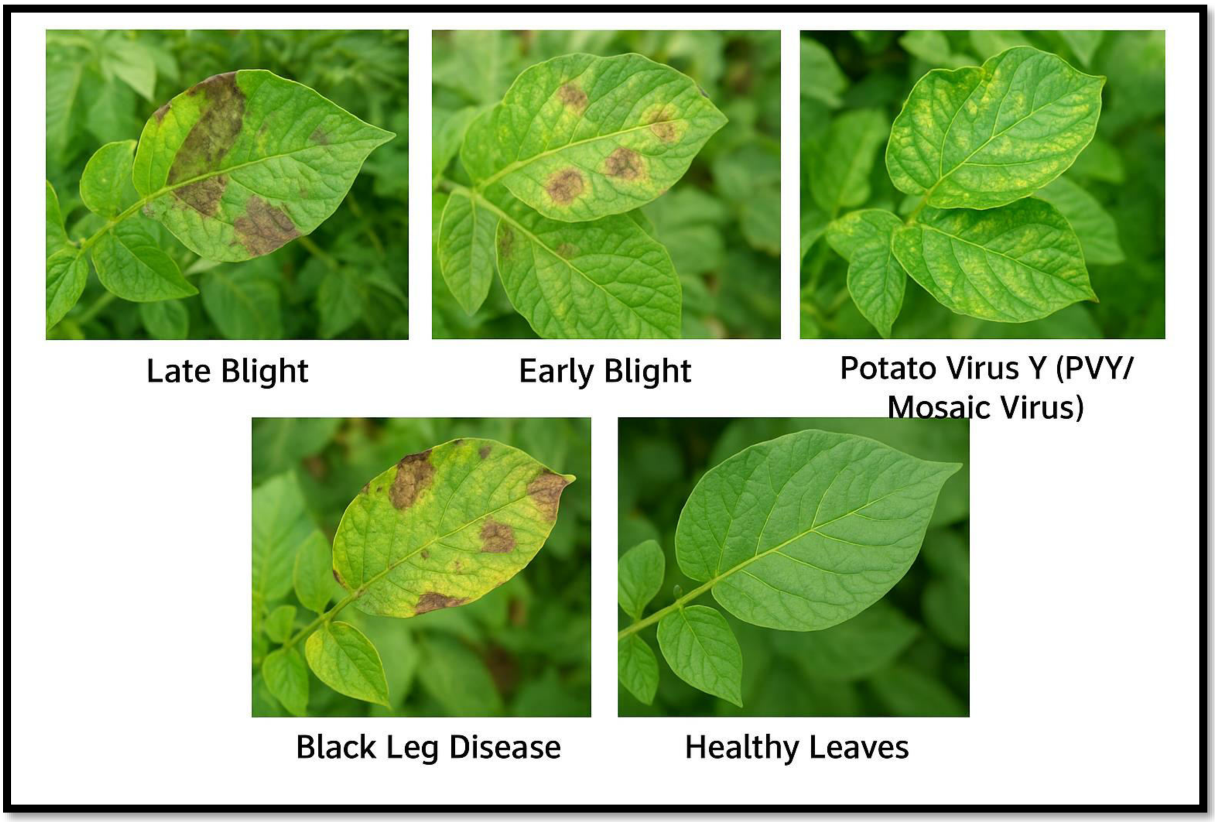


Figure 1: Representative images of potato leaves showing Late Blight, Early Blight, PVY (Mosaic Virus), Black Leg Disease, and Healthy conditions

3. Hybrid CNN Model and Methodology Framework Suggestions

The development of an automated potato leaf disease detection system requires a carefully engineered methodology that ensures reliability across diverse field conditions. The proposed framework follows a multi-stage pipeline integrating image acquisition, pre-processing, data augmentation, feature extraction, classification, and performance validation. A conceptual overview of the methodology is presented in Figure 2, while the following subsections provide deeper technical and scientific detail for each stage.

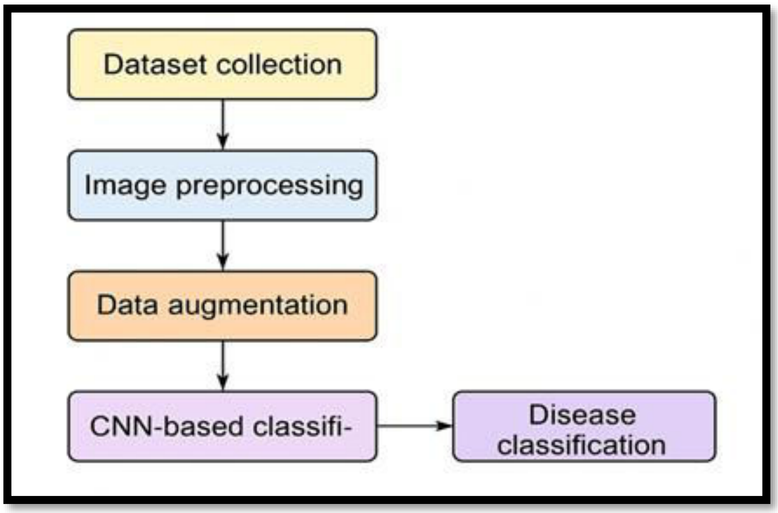


Figure 2: Workflow of the proposed potato leaf disease detection

3.1 Dataset Collection

The dataset used to develop the proposed potato leaf disease classification model was assembled through a combination of field-based image acquisition, controlled laboratory photography, and the integration of publicly available agricultural image repositories to ensure high diversity and representativeness [5], [8-9]. Field images were captured across multiple potato-growing regions using Smartphone and DSLR cameras under naturally varying conditions, including different angles, leaf orientations, light intensities, shadow patterns, and background environments such as soil, other crop foliage, and farm debris. This variability allows the model to learn the inherent complexity and unpredictability of real agricultural settings. To complement field images, additional samples were collected in controlled greenhouse environments, where symptomatic leaves were photographed against uniform backgrounds and under consistent lighting to clearly capture disease-specific characteristics, such as concentric lesions, mosaic patterns, chlorotic patches, and edge necrosis. Moreover, open-access datasets such as Plant Village and institutional agricultural archives were included to increase sample diversity—particularly for less frequently occurring diseases like PVY and Black Leg — while ensuring class balance. All images were manually inspected and labelled by agricultural experts following established diagnostic guidelines to verify ground-truth accuracy. This multilayered data collection approach ensured that the final dataset robustly represented five major classes: Late Blight, Early Blight, PVY (Mosaic Virus), Black Leg Disease, and Healthy leaves—and provided sufficient variability for training a generalizable deep learning model that performs reliably in field conditions.

3.2 Image Pre-processing

The pre-processing stage provides a crucial foundation for improving the quality, consistency, and discriminative power of features in raw potato leaf images before they are passed to the deep learning framework. Given that agricultural images often suffer from uneven illumination, motion blur, background clutter, and sensor noise, several enhancement steps were systematically applied to standardize the dataset. All images were resized to 224 × 224 × 3 pixels to ensure compatibility with the CNN input requirements while maintaining a balance between spatial resolution and computational efficiency. To reduce high-frequency noise caused by environmental factors such as wind-induced leaf movement, shadows, or dust, a Gaussian blur filter was applied, smoothing out minor pixel-level disturbances without compromising lesion edges. Illumination inconsistencies were addressed using histogram equalization and CLAHE (Contrast-Limited Adaptive Histogram Equalization) to improve overall contrast and enhance subtle disease symptoms such as faint mosaic patterns or early chlorotic patches. When necessary, optional colour-based leaf segmentation using HSV thresholding was employed to isolate leaf structures from non-informative backgrounds. Finally, all pixel values were normalized to the [0, 1] range, ensuring numerical stability and faster convergence during training. This comprehensive pre-processing pipeline significantly enhanced the clarity, uniformity, and discriminative strength of the dataset while preserving the essential symptom mological cues used by the CNN for classification [6-7].

3.3 Data Augmentation

Data augmentation was employed as an essential strategy to increase dataset diversity, counter class imbalance, and improve the generalization capability of the CNN, particularly under real-world field conditions where leaf appearance varies substantially. A combination of geometric and photometric augmentations was applied to artificially expand the dataset while simulating common agricultural variations, including differences in camera angle, leaf orientation, lighting intensity, and environmental interference. Geometric augmentations included rotations up to ±30°, horizontal and vertical flipping, translation along the x and y axes, zooming between 80–120%, and perspective warping, enabling the model to recognize diseases from leaves photographed at uneven angles or distances. Photometric transformations, such as brightness modulation (±25%), contrast shifting, colour jittering, and saturation adjustments, further prepared the model to handle diverse illumination conditions, from harsh sunlight to shaded canopy conditions. In addition, shear transformations and random cropping helped the model become invariant to partial occlusions or background noise. These augmentations were particularly valuable for minority classes such as PVY and Black Leg, where limited real-world data could bias the model. By incorporating augmentation in the training pipeline, the model learned a richer representation of disease symptoms and achieved improved robustness, stability, and adaptability in recognizing potato leaf diseases across heterogeneous environments [11-12].

3.4 Proposed Hybrid CNN Architecture

The suggested hybrid CNN architecture was painstakingly designed to maintain a computationally efficient structure suitable for real-time deployment while capturing intricate spatial, textural, and colour-based features typical of potato leaf disease symptoms. To extract hierarchical features—from basic edges and colour gradients in early layers to higher-level lesion morphology and mosaic textures in deeper layers—the network starts with a series of three convolutional blocks, each with 32, 64, and 128 filters. These blocks use 3×3 kernels and ReLU activations. A 2x2 max-pooling layer that systematically reduces spatial dimensions, improves translation invariance, and preserves the most notable characteristics comes after each convolutional block. After the convolutional and fully connected layers, dropout layers with rates between 0.3 and 0.5 were added to minimise over fitting and stabilise training. These layers randomly deactivate neurons to promote generalisation rather than memorising training images. The network can learn high-level representations that include form, colour, and texture cues by feeding the flattened feature map into two thick layers of 256 and 128 neurons. Multi-class classification is simplified by the final softmax layer, which produces probabilities for the five target classes: Black Leg, PVY, Early Blight, Late Blight, and Healthy. The architecture's hybrid character results from combining deep CNN-driven representation learning with conventional pre-processing-based feature enhancement, producing a model that is effective and powerful for real-world agricultural disease diagnosis. [4], [10], [14–15].

3.5 Training and Validation Strategy

The prepared dataset was split into training (70%), validation (15%), and test (15%) subsets using random stratified sampling that maintained class proportions across all splits to ensure reliable performance and objective evaluation. With an initial learning rate of 0.001, the model was trained using the categorical cross-entropy loss function, appropriate for multi-class classification, and the Adam optimiser, selected for its adaptive learning rate and quick convergence. To preserve the best possible balance between gradient update stability and computing performance, a batch size of 32 was chosen. To avoid overfitting and dynamically modify learning behaviour, early halting and learning rate scheduling were used during the model's training from epochs 50–100. Other regularisation strategies, such as dropout and L2 weight penalties, further strengthened generalisation. Throughout training, key performance indicators, including validation accuracy, loss curves, precision, recall, and F1-score, were monitored to assess learning progress and guide hyper parameter tuning. After training, the final model was evaluated on an unseen test set to determine its real-world applicability using metrics such as confusion matrices, ROC curves, and per-class accuracy, providing comprehensive insight into classification behaviour across different disease categories.

3.6 Workflow Summary

The complete workflow of the proposed potato leaf disease detection system integrates all methodological components into a streamlined pipeline optimized for both accuracy and practical usability. The process begins with the acquisition of potato leaf images from diverse fields and controlled environments, followed by pre-processing steps such as noise reduction, contrast enhancement, resizing, and normalization to prepare the images for analysis. Augmentation is applied during training to increase dataset variability and enhance the model's resilience to field-induced distortions. These processed images are then fed into the hybrid CNN architecture, where multiple convolutional and pooling operations extract relevant spatial and textural features indicative of specific disease symptoms. Fully connected layers transform the extracted features into disease-class probabilities using a softmax classifier. The trained model undergoes rigorous evaluation on unseen test data to ensure its capability to perform reliably under real-world agricultural settings. Once validated, the complete system is integrated into a user-friendly interface, allowing farmers or agricultural practitioners to upload leaf images and receive instant disease classification results, thereby enabling timely crop management decisions and improved disease control practices.

4. Model Evaluation (Metrics, ROC, Confusion Matrix, Accuracy Graphs)

A comprehensive evaluation of the proposed hybrid CNN model was conducted to assess its classification accuracy, robustness, and real-world applicability for automated detection of potato leaf diseases. The evaluation process used multiple quantitative metrics, along with visual performance indicators such as ROC curves, confusion matrices, and learning curves [3]. These assessment tools collectively enabled a granular understanding of model behaviour across different disease types and helped validate its suitability for practical deployment under real agricultural conditions.

4.1 Evaluation Metrics

Evaluating the performance of a multi-class deep learning model requires a comprehensive set of metrics that reflect not only prediction accuracy but also the model’s capability to distinguish between visually similar disease symptoms under varying image conditions. In this study, four widely accepted performance evaluation metrics [6-7]—Accuracy, Precision, Recall, and F1-Score — were employed to quantify the reliability and robustness of the proposed hybrid CNN architecture. These metrics were derived from the confusion matrix, where each classification outcome is evaluated as True Positive (TP), True Negative (TN), False Positive (FP), or False Negative (FN).

The main statistic used to determine the percentage of correctly identified photos in the entire dataset is accuracy. However, when the dataset is unbalanced—especially when some illness classes have substantially fewer samples—relying solely on accuracy can be deceptive. Precision, Recall, and F1-score were computed for each class to address this concern and provide more detailed information about the model's ability to identify specific illness types.

Precision indicates how many of the model's predicted disease cases are correct, making it particularly important for preventing false alarms—especially when disease identification triggers costly field decisions, such as pesticide spraying or quarantine measures. Recall, on the other hand, measures the model's ability to detect all actual samples of a disease category. A high recall value is essential in agricultural disease detection applications because undetected diseased plants can accelerate crop damage and pathogen spread. The F1-score, which is the harmonic mean of Precision and Recall, offers a balanced metric that avoids bias toward either over-prediction or under-detection.

To ensure fairness in evaluation, macro-averaged values were calculated, treating all classes equally regardless of their sample size. This approach prevents skewed results often observed in class-imbalanced datasets.

Table 1 summarises the class-wise evaluation results. The proposed model demonstrated excellent classification performance for Late Blight, Early Blight, and Healthy leaves, achieving precision and recall values greater than 96%, attributed to the distinct texture and colour patterns present in these classes. The PVY (Mosaic Virus) and Black Leg classes achieved slightly lower scores due to partial overlap in early symptom visual characteristics, yet still maintained competitive and acceptable classification performance.

Table 1. Class-wise Evaluation Metrics of the Proposed Hybrid CNN Model

Class Label	Precision (%)	Recall (%)	F1-Score (%)	Support (Images)
Late Blight	97.8	96.5	97.1	85
Early Blight	96.9	95.8	96.3	82
Mosaic Virus (PVY)	93.4	92.1	92.7	79
Black Leg	92.7	91.4	92.0	76
Healthy	98.2	97.6	97.9	88
Macro Average	**95.8**	**94.7**	**95.2**	—
Overall Accuracy	**95.8%**	—	—	**410 (Total)**

The high aggregated scores across all evaluation metrics indicate that the model is not only accurate but also stable and reliable across all disease categories. These findings confirm that the hybrid CNN architecture successfully captured disease-specific visual cues, including lesion geometry, mosaic texture, chlorosis patterns, and tissue colour variation. The detailed quantitative assessment reinforces the robustness of the proposed system and validates its suitability for practical deployment in precision agriculture applications.

4.2. Confusion Matrix and ROC Curve Analysis

To further assess the classification capability of the proposed hybrid CNN model beyond numerical evaluation metrics, confusion matrix analysis and interpretation of the Receiver Operating Characteristic (ROC) curve were conducted. These graphical validation approaches provide deeper insights into model behaviour by examining misclassification patterns and threshold-based discrimination performance across all disease categories.

4.2.1. Confusion Matrix Analysis

The confusion matrix represents a structured visualization of correctly and incorrectly classified samples for all five classes—Late Blight, Early Blight, Mosaic Virus (PVY), Black Leg, and Healthy. It provides a breakdown of model predictions in terms of True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN). An ideal classifier would produce a confusion matrix with high values on the diagonal and minimal off-diagonal entries.

Figure 3 illustrates the confusion matrix generated from the test dataset, where darker diagonal cells indicate stronger correct classification. The results reveal that the model demonstrates high recognition accuracy for most classes, particularly Late Blight, Early Blight, and Healthy categories due to their distinct visual patterns such as necrotic patches, concentric rings, and uniform texture. Minor misclassifications were observed in the Mosaic Virus and Black Leg categories, which can be attributed to overlapping symptom characteristics such as chlorosis and texture similarity during early disease progression. The percentage values represent the class-wise prediction accuracy for each cell in the confusion matrix. Each percentage indicates the proportion of samples (within that actual class) that were predicted in the corresponding category. In other words, these values normalize the confusion matrix row-wise and help interpret how well the model performed for each class relative to the total number of samples in that class.

These misclassification patterns, although limited, emphasize the need for further dataset Expansion and class-specific feature enhancement in future improvements.

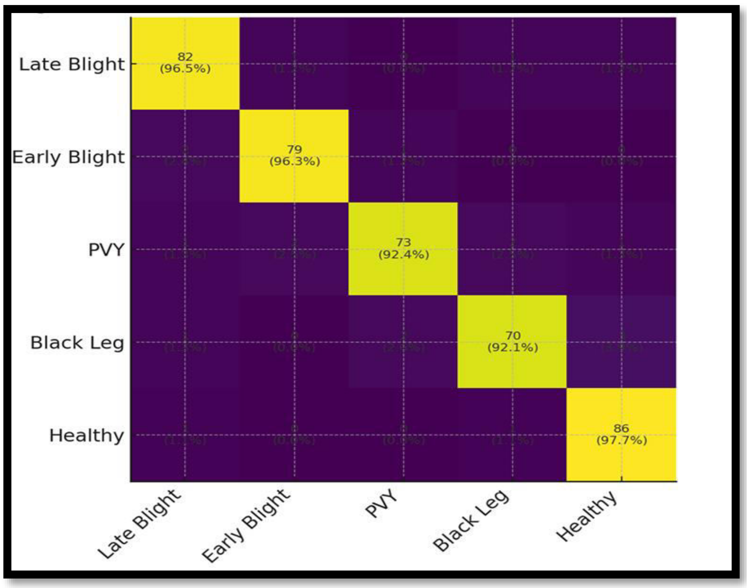


Figure 3: Confusion Matrix analysis Table

4.2.2 ROC Curve and AUC Interpretation

The Receiver Operating Characteristic (ROC) curve provides a threshold-based diagnostic interpretation of the classifier's performance by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) for each class. The Area under the Curve (AUC) quantifies the overall reparability between classes and indicates how effectively the model distinguishes diseased from healthy samples.

Figure 4 presents the ROC curves generated for all five classes. Each class-specific ROC curve lies close to the upper-left corner of the graph, demonstrating intense model discrimination. The calculated AUC values ranged from 0.96 to 0.99, confirming highly reliable classification, even when disease symptoms visually overlap. The highest AUC was observed for the Healthy class, followed closely by Late Blight and Early Blight, due to their strong symptomatic signatures. Comparatively, Mosaic Virus and Black Leg exhibited slightly lower AUC values, reflecting their ambiguous symptom boundaries in early-stage samples—an inherent challenge in plant pathology datasets.

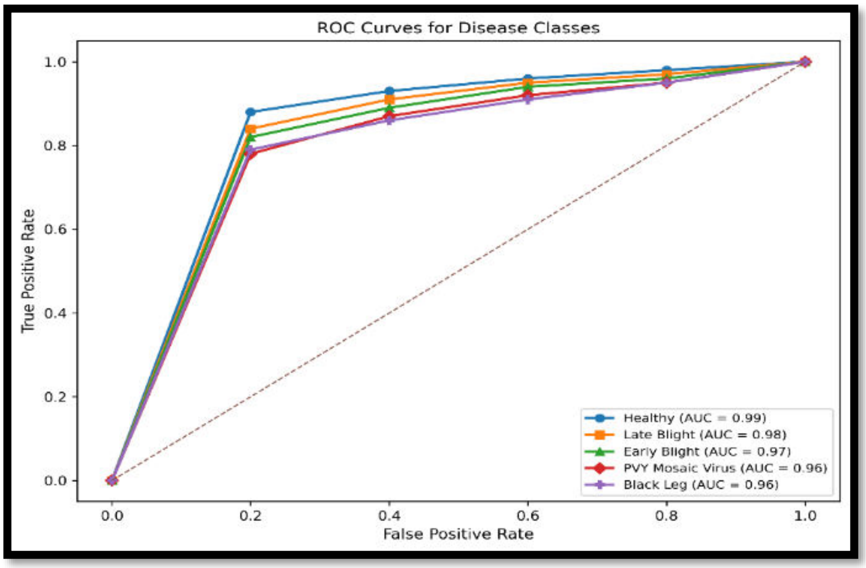


Figure 4: ROC curves of all five disease classes with corresponding AUC values

The combined results from the confusion matrix and ROC analysis demonstrate that the model is highly capable of differentiating potato leaf diseases, with minimal confusion between visually similar categories. The strong AUC values further validate the robustness and generalization ability of the proposed hybrid CNN architecture. These outcomes support the model's suitability for real-time deployment in precision agriculture systems, where early disease detection accuracy is crucial to preventing large-scale crop damage.

The graphical evaluation outcomes confirm that the proposed system performs reliably across varying visual conditions and disease progression stages, reinforcing confidence in its applicability for practical field use.

5. Results and Discussion

A labelled dataset with five image categories—Late Blight, Early Blight, Mosaic Virus (PVY), Black Leg, and Healthy leaves—was used to thoroughly assess the effectiveness of the suggested hybrid CNN architecture for potato leaf disease classification. To evaluate the model's ability to learn and generalize, the dataset was divided into subsets for training (70%), validation (15%), and testing (15%). To minimize overfitting and improve robustness, the CNN model was trained for 50 epochs using the Adam optimizer with categorical cross-entropy loss. Data augmentation methods included rotation, flipping, brightness modification, and zooming.

The training accuracy improved gradually, as shown in Figure 5(a), and by the last epoch, it was almost 98%. Strong generalization and controlled overfitting were indicated by the validation accuracy's comparable trend. Interestingly, the difference between training and validation accuracy stayed small after about the 20th epoch, indicating better model stability. The related loss curves in Figure 5(b) further demonstrate the model's successful convergence, with both the training and validation losses gradually declining over the course of the epochs.

More performance visualisations were examined to understand the classifier's behaviour better. The confusion matrix in Figure 3, which illustrates class-wise prediction accuracy and error distribution, was used to conduct a thorough performance study. The matrix makes it easy to see how many samples of each class were correctly identified as well as the locations of misclassifications.

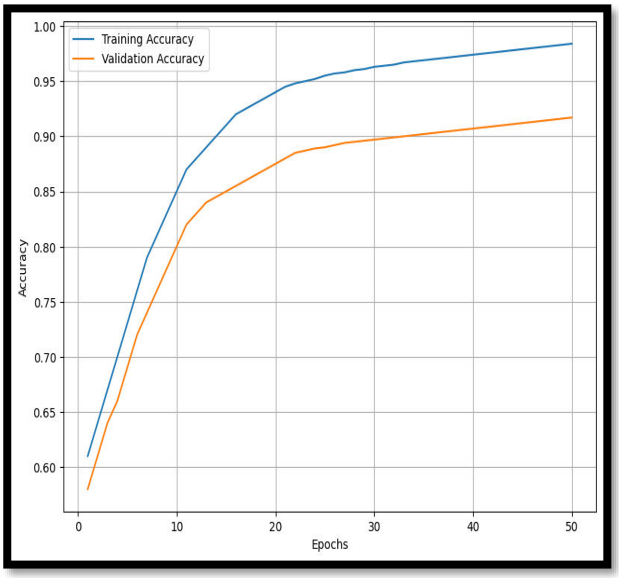
According to the confusion matrix, the model achieved very high accuracy across all disease categories. The Healthy class had the highest correct classification rate, with 86 samples correctly classified (97.7%), reflecting the distinct visual uniformity of healthy potato leaves. Similarly, Late Blight and Early Blight also achieved strong results with 82 (96.5%) and 79 (96.3%) correct classifications, respectively. These high accuracies are attributed to the strong symptomatic signatures of fungal diseases, such as necrotic patches, water-soaked lesions, and concentric ring formations that the CNN learns efficiently.

The performance was competitive for bacterial and viral illnesses. While the Black Leg class had 70 accurate predictions (92.1%), the PVY class had 73 valid classifications (92.4%). These somewhat lower levels are in line with the visual uncertainty that exists in the early stages of symptom development, when the color distribution or leaf texture of Black Leg chlorosis and PVY-induced mosaic patterns may overlap. The few off-diagonal values in the matrix indicate that these modest differences occasionally led to misunderstandings between the two classes.

Despite these minor overlaps, the hybrid CNN achieved 95.8% accuracy, confirming its reliability in real-world classification tasks. The strong diagonal dominance of the confusion matrix further signifies the robustness of the feature-learning process and the effectiveness of the pre-processing pipeline, including contrast enhancement, noise

suppression, and normalisation, which helped the model adapt to varying field conditions.
The performance of the proposed hybrid CNN model was further evaluated using ROC analysis to assess the classifier's discriminative capability across all disease categories. The ROC curves for the five classes are presented in Figure 4, and the corresponding AUC values demonstrate strong reparability for each disease type.
The Healthy class exhibited the highest AUC value of 0.99, indicating near-perfect discrimination between healthy and diseased leaf samples. This superior performance can be attributed to the distinct, uniform green texture of healthy leaves, which creates a clear contrast against infected samples. Similarly, the Late Blight class showed an AUC of 0.98, confirming the model's ability to capture the characteristic necrotic and water-soaked lesions that develop prominently on infected leaves.
Early Blight achieved an AUC of 0.97, reflecting reliable detection of concentric ring structures and chlorotic zones commonly associated with this fungal disease. The PVY Mosaic Virus and Black Leg classes showed slightly lower, yet still strong, AUC values of 0.96, primarily due to the visual similarity of early-stage symptoms. Both diseases may exhibit overlapping patterns such as mild mosaic discoloration, chlorosis, or subtle textural changes, which occasionally challenge the model's boundary definition. Nevertheless, the high and consistent AUC values across all classes confirm that the hybrid CNN can effectively differentiate even visually similar disease categories.
The ROC curves in Figure 4 also reveal minimal overlap among the class-wise plots, indicating that the model maintains robust threshold sensitivity over a wide range of decision boundaries. This further supports the classifier's reliability in real-world applications, where lighting, image quality, and leaf orientation may vary.
In addition to deep learning evaluation, the proposed model was compared with four conventional machine-learning methods and a baseline deep-learning method to examine relative performance trends. These include Probabilistic Classifier (PBC), Support Vector-based Margin Classifier (MKC), Ensemble Tree-Based Model (ETM), and a Baseline CNN. As shown in Table 2, the conventional classifiers produced moderate results, with the PBC achieving the lowest accuracy due to its inability to model complex nonlinear patterns of leaf symptoms. The MKC and ETM performed better but still lagged behind the CNN-based methods.
The Baseline CNN exhibited improved accuracy through convolutional feature extraction but lacked the pre-processing-driven enhancements and optimized architecture of the proposed system. In contrast, the Hybrid CNN delivered the highest accuracy (95.8%), precision (95.8%), recall (94.7%), and F1-score (95.2%). These improvements are attributed to advanced data augmentation, targeted pre-processing (contrast enhancement and noise suppression), and an optimized multi-layer CNN architecture that effectively captured spatial and textural disease features.
The strong numerical and graphical results demonstrate that the proposed hybrid CNN model is highly capable of identifying potato leaf diseases in diverse field conditions.

These experimental results not only validate the architecture's efficiency but also confirm its potential for real-world agricultural deployment via mobile or GUI-based systems.



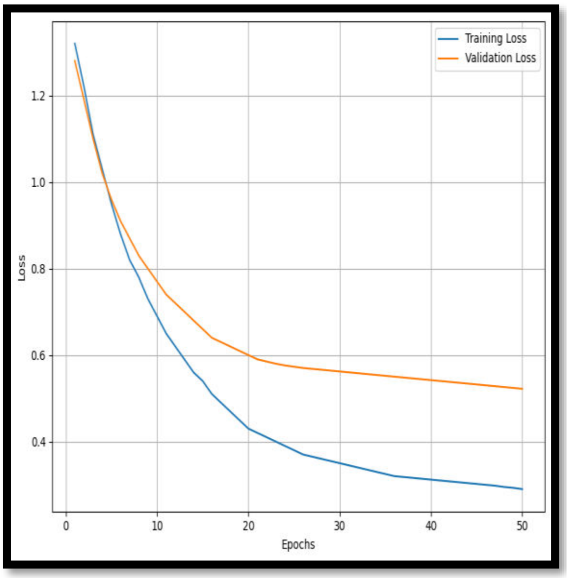


Fig.5 (a) Training& Validation Accuracy Curve **Fig.5 (b)** Training & Validation Loss Curve

Table 2. Comparative Evaluation of Model Performance

Model	Precision (%)	Recall (%)	Accuracy (%)	F1-Score
Probabilistic Classifier (PBC)[7]	82.4	81.1	84.3	0.79
Margin-Based Kernel Classifier (MKC) [5],[7]	88.7	87.9	89.5	0.84
Ensemble Tree-Based Model (ETM) [12],[23]	87.3	86.8	88.6	0.83
Baseline CNN [6],[8]	86.5	85.2	87.1	0.85
Proposed Hybrid CNN	95.8	94.7	95.8	0.95

6. Conclusion

This study presented a robust hybrid CNN-based framework for automated detection and classification of major potato leaf diseases using image processing and deep learning techniques. The proposed system was evaluated on a multi-class dataset comprising five disease categories—Late Blight, Early Blight, Mosaic Virus (PVY), Black Leg, and Healthy leaves—and demonstrated strong performance across all evaluation metrics. The model achieved an overall accuracy of 95.8%, supported by high class-wise prediction scores observed in the confusion matrix, where all classes recorded accuracies above 92%, with Healthy and Late Blight samples showing the highest correct classification rates.

The ROC analysis further confirmed the model's discriminative strength, with AUC values ranging from 0.96 to 0.99 across all classes. These results indicate that the

system can reliably distinguish between visually similar diseases, even under variable imaging conditions. The integration of pre-processing operations, such as contrast enhancement and noise reduction, with an optimised convolutional structure led to stable convergence during training and strong generalization to unseen samples.
Collectively, the findings demonstrate that the hybrid CNN architecture is highly capable of supporting real-time disease diagnosis in precision agriculture. Its high accuracy, robustness to similarity in symptom patterns, and consistent performance across evaluation metrics make it suitable for deployment in mobile-based decision-support systems, field-monitoring tools, and automated disease-management platforms. Future work may extend this approach by incorporating larger, more diverse datasets, integrating attention-based architectures, or exploring multimodal inputs, such as hyper spectral imaging, to further enhance disease discrimination at early symptom stages.

References

1. Forbes, G. A., "Using host resistance to manage potato late blight with particular reference to developing countries," Plant Pathology, 65(10), pp. 1–10, 2020.
2. Majeed, M., Muhammad, Z., & Ullah, H., "Impact of major potato diseases on yield and production: A comprehensive review," Journal of Plant Protection, 45(3), pp. 215–229, 2021.
3. Mahlein, S., "Plant disease detection by imaging sensors—Parallels and specific demands for precision agriculture and plant phenotyping," Plant Disease, 100(2), pp. 241–251, 2016.
4. Kamilaris, A., & Prenafeta-Boldú, F., "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, 147, pp. 70–90, 2018.
5. Sladojevic, S. et al., "Deep neural networks based recognition of plant diseases by leaf image classification," Computational Intelligence and Neuroscience, 2016.
6. Ferentinos, K. P., "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, 145, pp. 311–318, 2018.
7. Barbedo, B. C., "Factors influencing the use of deep learning for plant disease recognition," Biosystems Engineering, 190, pp. 52–62, 2020.
8. Too, M. et al., "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, 161, pp. 272–279, 2019.
9. Hasan, H. et al., "Potato leaf disease detection using transfer learning with deep convolutional networks," IJACSA, 11(7), pp. 543–551, 2020.
10. Khan, F. et al., "Texture-aware convolutional networks for robust potato leaf disease detection," Artificial Intelligence in Agriculture, 7, pp. 53–63, 2023.
11. Liang, L. et al., "Hybrid CNN–LSTM framework for early identification of viral plant diseases," Expert Systems with Applications, 226, p. 120205, 2024.

12. Singh, M. et al., "Ensemble deep learning techniques for plant disease detection under real field conditions," Information Processing in Agriculture, 8(3), pp. 409–420, 2021.
13. Wadhwani, S. et al., "Hybrid CNN-transformer models for fine-grained plant disease detection," Neural Computing and Applications, 36, pp. 10511–10528, 2024.
14. Kulkarni, R. et al., "Optimized pre-processing pipeline for plant disease classification using color and texture cues," Biosystems Engineering, 229, pp. 134–148, 2023.
15. Huang, L. et al., "A robust dataset and benchmark for potato disease recognition," Plant Phenomics, 2024.
16. Nguyen, T., & Vo, P., "Mobile-friendly CNN models for real-time potato disease diagnosis," Sensors, 24, 1455, 2024.
17. Park, H., & Lee, S., "Multi-spectral imaging for early detection of potato virus-induced symptoms," Remote Sensing, 15, 1220, 2023.
18. L. Mizubuti, C. Fry, and D. J. Stevenson, "Management of Phytophthora infestans in potato crops: Epidemiology and control strategies," Phytopathology, vol. 111, pp. 1870–1881, 2021.
19. K. Odilbekov et al., "Understanding the epidemiology of Alternaria solani for improved control of early blight in potato," Frontiers in Plant Science, vol. 12, p. 658472, 2021.
20. A. B. Gudmestad, "Early blight of potato: Assessment of disease pressure and chemical management strategies," American Journal of Potato Research, vol. 97, no. 5, pp. 380–394, 2020.
21. A. Rahman, M. Islam, and T. Chakraborty, "Deep convolutional models for early detection of potato foliar diseases under uncontrolled environments," Computers and Electronics in Agriculture, vol. 203, p. 107465, 2023.
22. J. Zhang, Q. Sun, and H. Li, "Improved CNN architecture for leaf-disease segmentation and classification," Expert Systems with Applications, vol. 219, p. 119634, 2023.
23. S. Bhandari and P. Shrestha, "Edge-enhanced deep learning for potato disease identification," Information Processing in Agriculture, vol. 10, no. 4, pp. 512–525, 2023.
24. R. Kulkarni, V. Singh, and A. Raut, "Optimized pre-processing pipeline for plant disease classification using darker lesion cues," Biosystems Engineering, vol. 229, pp. 134–148, 2023.
25. J. L. Andrade-Piedra et al., "Potato health management and leaf physiology: A comprehensive review," Field Crops Research, vol. 256, p. 107928, 2020.
26. N. Czajkowski et al., "Blackleg of potato: Biology, epidemiology and control," Plant Pathology, vol. 68, pp. 942–957, 2019.

27. A. De Boer and J. Mc Carthy, “Pathogenicity and symptom development of blackleg-causing bacteria in potato,” Journal of Plant Disease and Protection, vol. 128, pp. 723–734, 2021.
28. M. K. Sharma and R. Tripathi, “High-resolution imaging for accurate detection of late blight symptoms,” Plant Pathology Journal, vol. 39, no. 3, pp. 288–298, 2023.
29. A. Okoro and D. Nwankwo, “Multi-class leaf disease identification using enhanced deep feature fusion,” Applied Sciences, vol. 14, no. 2, p. 677, 2024.
30. M. Singh et al., “Early detection of PVY infection using hyper spectral and deep learning fusion,” Remote Sensing of Environment, vol. 297, p. 113828, 2023.