Potato J (2025) 52 (2): 183-196

ENHANCING POTATO CROP HEALTH MONITORING USING TWO STAGE CNN BASED DISEASE CLASSIFICATION MODELS

Aditya A. Shastri[1], Astha More[1], Tisha Soni[1], Milind Ratnaparkhe[2], Sashi Rawat[3], Ketan Sabale[1*]and Manish Paliwal[1]

ABSTRACT: Potato is one of the most extensively cultivated crops in India as well as worldwide and serves as a staple food in many regions. Due to its agricultural and economic importance, effective disease management is essential to ensure healthy crop yields. Traditional methods of disease detection rely on manual visual inspection by farmers or agricultural experts, which often lack precision and are prone to misdiagnosis, leading to substantial crop losses. This study proposes a deep learning-based approach for automated disease detection in potato plants using a Convolutional Neural Network (CNN). The novelty of this work lies in developing a two-stage CNN framework that first classifies the potato leaf images as either healthy or unhealthy. Next, the unhealthy leaf images are further classified as either early blight or late blight diseases. The proposed CNN models are optimized using the Adam optimizer with a learning rate of 0.0001. Extensive experimentation on a dataset of 1,500 images demonstrates high classification accuracies of 98.3% and 99% for the two stages, respectively. The results demonstrate that our proposed CNN models are highly effective for automated disease detection and improve decision-making in potato crops.

KEYWORDS: Potato Disease, Convolutional Neural Network (CNN), Deep Learning in Agriculture, Early Blight, Late Blight.

INTRODUCTION

Agriculture is one of the most important aspects of modern-day society as it is used for food, economic development, and global food security. Potato is an important food in human cuisine, which is frequently consumed staple foods. It is the third most important food crop in the world after wheat and rice, as per the Food and Agriculture Organization (FAO), a production of more than 368 million metric tons in the year 2021 (Rashid, *et al.*, 2021). This plant is a combination of carbohydrate, vitamins, and minerals for hundreds of millions of people. However, potatoes are very vulnerable to disease, i.e., fungal infections like early blight and late blight, that cause unprecedented loss of production.

The development of disease to plant undermines food security, adds to the production cost, and causes financial losses for farmers. Hence, early detection of potato leaf disease is of utmost importance in avoiding such threats and minimize dependence on extensive usage of pesticides (Strange & Scott, 2005). Traditional methods

---

*Corresponding author; email: ketan.sabale@sot.pdpu.ac.in

[1]Department of Computer Science and Engineering, School of Technology, Pandit Deendayal Energy University, Gandhinagar, India.

[2]ICAR – Indian Institute of Soybean Research, Indore, India.

[3]ICAR – Central Institute of Agricultural Engineering, Bhopal, India

involve manual observation by experts, which is time-consuming and prone to human error. Due to these shortcomings, computer vision and machine learning approaches have become useful for automation.

Traditional approaches extract features such as colour and texture from leaf and apply classifier to distinguish between the disease and non-disease plants (Dalal & Triggs, 2005). Other approaches such as Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP) have also been used for leaf disease detection with better accuracy (Pujari, Yakkundimath, & Byadgi, 2015). However, with the recent arrival of neural networks, deep learning-based approaches have emerged as a powerful tool for plant disease classification that outperforms traditional approaches by learning more complex patterns from images (Singh & Misra, 2017).

Recent studies have demonstrated the potential of CNNs in the classification of plant diseases. Hence, in this paper, we propose a two stage CNN-based framework for the detection of potato leaf diseases. In the first stage, we classify the leaf image as healthy or unhealthy. Further, the unhealthy image is again classified as early blight or late blight. This hierarchical design simplifies the classification task at each stage and improves the overall model performance. The advantage of using CNN is that it extracts features automatically from raw images and is therefore less sensitive to light variations, noise in the background, and variations in disease patterns (Kamilaris & Prenafeta-Boldú, 2018). It is difficult to distinguish early blight and late blight in their early stage using conventional methods as they exhibit similar texture and patterns. Both diseases cause leaf discoloration, necrotic lesions, and chlorosis, making visual diagnosis difficult (Mohanty, Hughes, & Salathé, 2016). Thus, the novelty of this work lies in developing a 3 convolutional layers CNN model that distinguishes early and late blight images accurately in its early stage.

We propose same model for classification at both the stages. Each model consists of three convolutional layers, incorporated with max pooling layers, followed by a flattening layer, a fully connected dense layer, and an output layer. A dropout layer with a dropout rate of 0.5 is used after the dense layer to avoid overfitting (Wahabzada *et al.*, 2015). The input images are resized to and normalized in [0, 1] range to ensure training uniformity. A series of experiments are conducted to determine the optimal values for hyperparameters (Chollet, 2017). These are presented in detail in the further sections.

The proposed two-stage CNN model is extensively evaluated on 1,500 potato leaf images using accuracy, precision, recall, and F1-score, that are obtained from the confusion matrices of both stages. In stage 1 (i.e., healthy vs. unhealthy classification), the model achieves an accuracy of 98.3% with minimal misclassifications. Similarly, in stage 2 (i.e., early vs. late blight classification), it achieves an accuracy of 99%. Apart from these, we test on various configurations to decide the optimal values of hyperparameters. The Adam optimizer with a learning rate of 0.0001 and three convolutional layers per stage give the best results. Finally, we obtain the training and validation curves to confirm early convergence without overfitting.

The rest of the paper is structured as follows. The Literature Review section provides a brief summary of previous studies on potato leaf identification. The proposed methodology and the architecture of our two -stage CNN model is discussed in Materials and Methods section. Results and Discussion section presents the experimental results and its accompanying discussions. Finally, conclusions and future work are provided in the last section.

LITERATURE REVIEW

In recent years, deep learning has emerged as an important and useful tool for automating disease diagnosis from plant leaf images. It offers high accuracy and scalability as compared to the traditional methods. Thus, in this section, we present a concise review of recent works that focus on potato leaf disease detection using convolutional neural networks (CNNs), transfer learning, hybrid architectures, explainable AI techniques etc.

One prominent approach integrates depthwise separable convolutions and attention mechanisms. The MDSCIRNet model is enhanced with multi-head attention and tested against both traditional deep learning and ensemble classifiers and achieved up to 99% accuracy when combined with an SVM classifier (Reis & Turk, 2024). Similarly, explainable AI was applied to enhance trust in prediction outputs. One such study employ LIME and SHAP interpretations to ensemble models like CNN, CNN-SVM, DNN. They achieved 99% accuracy while also improving transparency in decision-making (Paul *et al.*, 2024).

Lightweight CNNs like RegNetY-400MF have also shown promise in practical deployments. This model efficiently classified seven types of potato diseases with 90.68% accuracy, supporting real-time use on edge devices (Chang & Lai, 2024). There are many works that have focused on optimizing existing CNNs. An improved version of VGG16 (VGG16S) reduced parameters to one-tenth while achieving 97.87% accuracy (Zhang *et al.*, 2025). Additionally, an ensemble model based on ResNet50V2 and DenseNet201 accurately assessed disease stages beyond classification (Chowdhury *et al.*, 2024). Another modified version of VGG19, NASNetMobile, and DenseNet169, improved performance by adding additional layers and feature reductions. They demonstrated up to 99% accuracy and high AUC-ROC scores (Lanjewar, Morajkar & Payaswini, 2024).

Custom based CNN architectures have also been proven effective. A new CNN model designed specifically for potato leaves achieved a notable 98.28% accuracy, significantly outperforming prior models (Sofuoğlu & Bırant, 2024). Another work combined MDSCIRNet with SEResNet101V2 to attain 99.67% test accuracy (Bajpai, Sahu, & Tiwari, 2025). Furthermore, vision transformer (ViT_B_16) models have also been effective, with 99% accuracy on a combined dataset of early blight, late blight, and healthy samples (Adhikari, 2024).

One study highlighted the need for more robust models as VGG19 fail to distinguish healthy leaves (Fuadi, Putri, Nasien, & Oktarina, 2024). The EfficientNetB0 architecture was also tested on a potato leaf dataset, achieving 99.05% accuracy using compound scaling strategies (Upadhyay, Jain, & Prasad, 2024). While most studies focus solely on image-based features, some integrate environmental data. One such work used meteorological features selection algorithms like bGGO to boost the predictive power of traditional ML models, achieving 98.3% accuracy (Radwan, Alhussan, Ibrahim, & Tawfeek, 2024).

In terms of reviews, one comprehensive survey emphasized the importance of CNNs (like ResNet, MobileNet, VGG) in potato disease detection. It also discussed challenges such as generalizability and data scarcity (Gülmez, 2024). Finally, a hybrid deep learning model based on DenseNet-121 and Gaussian filter fusion produced a training accuracy of 99.08% and validation accuracy of 98.37% (Raza, Pitafi, Shaikh, & Ahmed, 2025).

These studies demonstrate that deep learning models are very popular in disease classification. They can produce good results if properly optimized with transfer learning, lightweight architectures, and ensemble strategies. The summary of a few works is presented in the Table 1 below.

MATERIALS AND METHODS

In this section, we present a detailed overview of the Convolutional Neural Network (CNN)-based model for potato leaf disease classification. This includes dataset description, preprocessing techniques applied to enhance the quality of images, and the design of a robust CNN architecture specifically designed for leaf disease classification. We also present the justification of each architectural choices, parameter tuning strategies, and training configurations used to optimize model performance. This flowgraph is presented in Fig. 1 below.

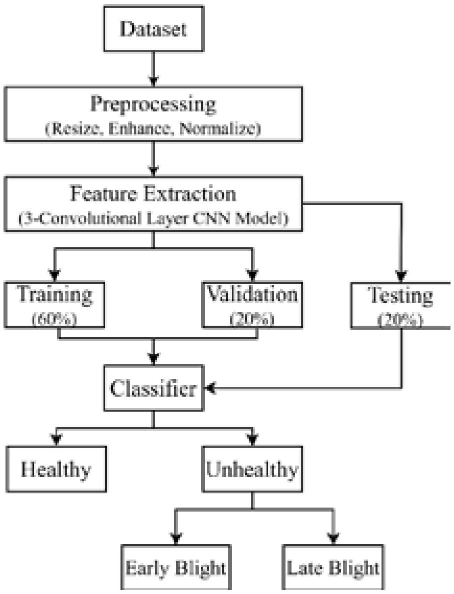


Fig. 1. Flowgraph of proposed two stage CNN based disease classification system.

Dataset Description

The dataset used in this study is taken from Kaggle, which is the standard public

Table 1.

Sr. No.	Authors	Work Done	Proposed Model	Accuracy
1	(Zhang, et al., 2025)	Designed a full DL pipeline for diagnosing potato leaf diseases.	CNN	97.87%
2	(Bajpai, Sahu, & Tiwari, 2025)	Integrated attention mechanisms and squeeze-and-excitation blocks to enhance detection accuracy.	CNN + SE + Attention	99%
3	(Adhikari, 2024)	Used Vision Transformer for classifying potato leaf diseases.	Vision Transformer	99%
4	(Chang & Lai, 2024)	Developed a lightweight CNN optimized for edge devices and mobile platforms.	Lightweight CNN	90.68%
5	(Chowdhury, et al., 2024)	Employed ensemble deep learning and classified severity levels of diseases.	Ensemble CNNs	Not mentioned
6	(Sofuo□lu & Bırant, 2024)	Applied standard DL approach for identifying various potato plant leaf diseases.	VGG16, ResNet50	98.2%
7	(Lanjewar, Morajkar, & Payaswini, 2024)	Introduced modified TL frameworks for better identification using pretrained models.	NasNet	95%
8	(Fuadi, Putri, Nasien, & Oktarina, 2024)	Evaluated various CNNs to find the most effective architecture.	AlexNet, VGG, ResNet	97%
9	(Upadhyay, Jain, & Prasad, 2024)	Targeted early and late blight detection using a pre-trained efficient model.	EfficientNetB0	99%
10	(Radwan, Alhussan, Ibrahim, & Tawfeek, 2024)	Used feature selection and ML optimization for disease classification.	SVM, RF + PCA/GA	98%

repository (Muhammad, 2025). This dataset consists of 1,500 labelled potato leaf images categorized into two main groups: 500 healthy and 1,000 unhealthy leaves. The unhealthy class is further divided into 500 images of Early Blight and 500 images of Late Blight. For this work, the dataset is split into three subsets for model training and evaluation: 60% of the images are used for training, 20% for testing, and the remaining 20% for validation. Table 2 below presents the distribution of samples across the categories and data splits. Furthermore, Fig. 2 below provides the sample images of Normal leaf, early blight leaf and late blight leaf.

Data Preprocessing

Data preprocessing is the most important step in classification model to ensure consistent input for the deep learning model and improve its performance. We employed three preprocessing steps to the input image as given below.

1. **Image Resizing**: All images were resized to pixels. This standardization makes all the images of same size, which is required for efficient processing by the CNN model. This also reduces computational overhead while preserving sufficient detail for accurate classification.

Table 2. Distribution of potato leaf images across health categories and dataset splits (*training, testing* and *validation*).

Category	Early Blight	Late Blight	Healthy	Total
Training	300	300	300	900
Testing	100	100	100	300
Validation	100	100	100	300
Total	500	500	500	1,500

(a) *(b)* *(c)*

Fig. 2. Representative sample images of potato leaves: (a) Healthy leaf, (b) Leaf affected by Early Blight, and (c) Leaf affected by Late Blight.

2. **Colour Image Enhancement**: Image enhancement techniques are required to improve the identification of features relevant to disease detection. There are many image enhancement methods available such as histogram equalization, contrast stretching, and CLAHE (Contrast Limited Adaptive Histogram Equalization). We use the CLAHE that is most commonly used for enhancement (Shastri, Tamrakar, & Ahuja, Density-wise two stage mammogram classification using texture exploiting descriptors, 2018).

3. **Normalization**: All pixel values are normalized in the range, which is the standard technique to accelerate the training process and contributes to the stability of gradient descent optimization by ensuring uniform input scales across the network. We use the following formula to perform normalization (Shastri, Tamrakar, & Ahuja, Density-wise two stage mammogram classification using texture exploiting descriptors, 2018):

$$I'\ (x, y) = \frac{I\ (x, y) - min\ (I)}{max\ (I) - min\ (I)'}$$

Where, is the pixel position, is the normalized pixel intensity, is the actual pixel intensity, is the minimum intensity over all the pixels, and is the maximum intensity over all the pixels.

Model Preparation

In this subsection, we present the architecture of our proposed model. As mentioned earlier, we address two binary classification problems using CNN. First, identifying potato leaves as healthy or unhealthy, and next, further classifying the unhealthy leaves as early blight or late blight. Since both tasks rely on the same type of image input (i.e., potato leaf images), we propose and implement a *single* CNN architecture for both the tasks. This unified architecture is designed in such a way that it is computationally efficient and capable of extracting high quality features capable of discriminating between the different disease classes. This is given in Fig. 2 below.

First, the proposed model's architecture begins with an input layer that accepts resized colour images of pixels. As mentioned earlier, this standardizes the input size and is useful in consistent data processing and efficient GPU utilization. The use of three channels (red, green, and blue) ensures that colour-based distinctions are retained in the analysis. This is important for identifying disease patterns such as yellowing, browning, or lesions.

Next, the architecture consists of a series of convolutional layers to extract features from the images. In our work, we use three different convolutional layers. The first convolutional layer employs 32 filters of size to capture low-level features such as edges and corners. The filter size of is selected for its efficiency in capturing local patterns without adding any extra computational cost. A second convolutional layer have 64 filters (again of size). This is introduced to enable learning of more complex features by using lower-level information. Finally, a third convolutional layer with 128 filters is introduced to capture high-level hidden features that mostly contributes for distinguishing between early and late blight symptoms. After every convolutional layer, we employ an activation function, specifically the Rectified Linear Unit (ReLU) function. ReLU is crucial for learning complex feature representations by introducing non-linearity to the model. It also improves computational efficiency and solves the vanishing gradient problem to accelerate convergence during training.

Furthermore, to reduce the dimensionality of features and retain only the most critical information, each convolutional block is followed by a max-pooling layer with a window size of. This step reduces computation and contributes to generalization by making the model invariant to small translations and distortions in the input images. The combination of convolution and pooling layers allows the model to progressively extract and condense important image features.

After the final convolution and pooling layers operations, we get a two-dimensional feature map. These maps are given as an input to a flatten layer that converts them into a one-dimensional vector. This transformation is necessary as the next layer is fully connected (dense) layer and it takes single vector as an input. Thus, this flattened vector is then passed into a dense layer containing 128 neurons. Here, 128 neurons are selected to avoid the risk of overfitting to ensure that the model remains generalizable on unseen data. This layer integrates information across the entire image and make informed predictions based on the learned features. We use a dropout rate of 0.5 to further reduce the chances of overfitting and improve generalization[#].

Finally, the output layer consists of a single neuron incorporated with a sigmoid activation function. This function is mostly suitable for binary classification tasks. It generates a probability value between 0 and 1, where a value close to 1 indicates the positive class (in our case, healthy in first

---

[#] Dropout rate of 0.5 randomly deactivates 50% of the neurons during training. This prevents the network from becoming overly dependent on specific path, thereby decreasing the chances of overfitting.

stage or early blight in second stage) and a value close to 0 indicates the negative class (again, unhealthy in first stage or late blight in second stage). A threshold of 0.5 is used for decision making. For example, if the predicted probability for a given input instance is less than 0.5, the instance is classified as belonging to the negative class; on the other hand, if the probability exceeds 0.5, it is assigned to the positive class. Fig. 3 below presents the architecture of proposed model with three convolutional layers, followed by flattening and fully connected layers.

The detailed technical overview of the proposed CNN model is given in Table 3 below. This table summarizes the architecture used in both stages of classification. The table includes following columns: the layer type, output shape, number of parameters, and the activation function applied to that layer. The output shape gives the dimensionality of the data and the activation functions indicate how non-linearity is introduced at each stage. This provides a clear understanding of the model's depth, complexity, and trainable capacity.

Table 3. Detailed architecture of the proposed CNN model (*for both classification stages*).

Layer Type	Output Shape	# of Parameters	Activation Function
Conv2D (32 filters,)	(126, 126, 32)	896	ReLU
MaxPooling2D ()	(63, 63, 32)	0	-
Conv2D (64 filters,)	(61, 61, 64)	18,496	ReLU
MaxPooling2D ()	(30, 30, 64)	0	-
Conv2D (128 filters,)	(28, 28, 128)	73,856	ReLU
MaxPooling2D ()	(14, 14, 128)	0	-
Flatten	(25088)	0	-
Dense (128)	(128)	3,211,392	ReLU
Dense (1)	(1)	129	Sigmoid
Total Parameters		33,04,769	

To find out the optimal architecture and the values of hyperparameters required for the CNN model, we perform a series of experimentations. For example, the models are trained using three optimizers (i.e., Adam, RMSprop, and SGD), and three different learning rates (0.001, 0.01, and 0.00005). Additionally, we explored architectures with 2, 3, and 4 convolutional layers for both classification stages. The configuration that produced the highest validation performance in each case was selected for final model training. This is discussed in detail in the next section.

RESULTS AND DISCUSSIONS

In this section, we present the experimental results and corresponding analysis for the

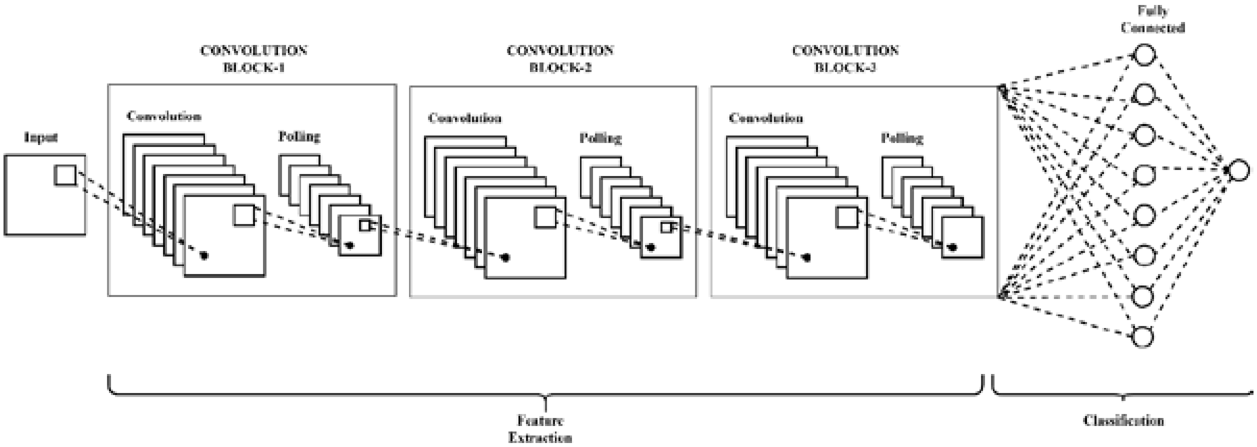


Fig. 3. Architecture of Proposed 3 convolutional layer CNN model.

proposed convolutional neural network (CNN)-based approach for classifying potato leaf diseases. As mentioned earlier, the evaluation is conducted in two stages: (1) classification of leaves as healthy or unhealthy, and (2) classification of unhealthy leaves as early blight or late blight. Furthermore, we also discuss the performance metrics, parameter tuning, and visualizations to validate the effectiveness of the proposed model.

First, to assess the classification performance, we compute four key evaluation metrics; namely accuracy, precision, recall, and F1-score as given by (Shastri, Prajapati, Katariya, Paliwal, & Sabale, 2025). These values are calculated based on the confusion matrix for each stage. All the values presented in the confusion matrix are obtained by experimenting on a held-out test set (20%) that is not used for training and validation. The confusion matrix for stage 1 is given in Fig. 4, and for stage 2 is given in Fig. 5 below. The confusion matrix gives us true positives (TP: healthy correctly predicted as healthy), true negatives (TN: unhealthy correctly predicted as unhealthy), false positives (FP: unhealthy misclassified as healthy), and false negatives (FN: healthy misclassified as unhealthy).

Using the values from the confusion matrix, the four metrics values are evaluated and given in the Table 4 below.

As evident from this table, the model performs exceptionally well for both stage classifications. Also, the relative low values

Table 4. Accuracy, Precision, Recall, and F1-Score values for both stages (healthy vs. unhealthy and early blight vs. late blight) classifications.

Evaluation Metrics	Formula	Healthy vs. Unhealthy	Early Blight vs. Late Blight
Accuracy	$Accuracy = TP + TN \mid TP + TN + FP + FN$	0.9833	0.9900
Precision	$Precision = TP \mid TP + FP$	0.9797	0.9900
Recall	$Recall = TP \mid TP + FP$	0.9700	0.9900
F1-score	$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$	0.9748	0.9900

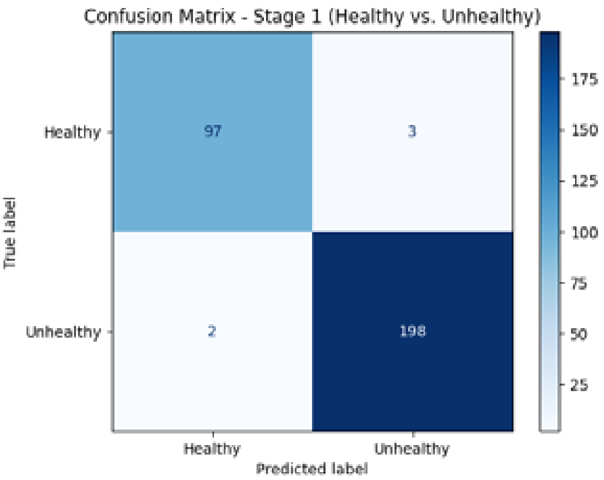


Fig. 4. Confusion matrix for Stage 1 (Healthy vs. Unhealthy)

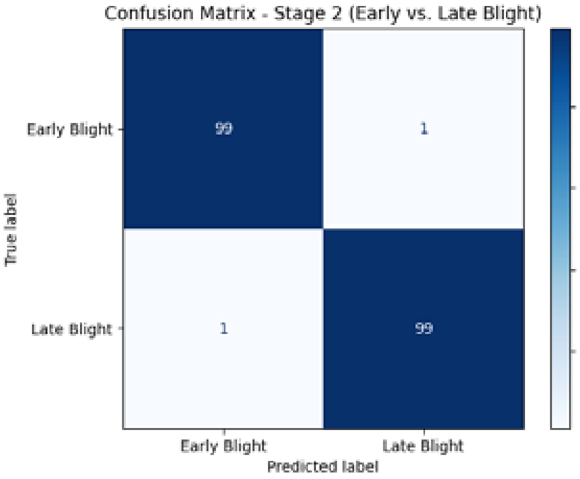


Fig. 5. Confusion matrix for Stage 2 (Early Blight vs. Late Blight)

of false positives and false negatives shows that the model is robust in discriminating the diseases.

Next, as mentioned earlier, to optimize the CNN model's performance, we perform an extensive hyperparameter tuning. For this we evaluate our model using three different optimizers; Adam, RMSprop, and SGD. Among these, Adam optimizer outperformed others in both stages of classification and hence, selected for the proposed model. This analysis is given in Fig. 6 below.

Following this, we investigate the effect of learning rate on the performance of our model. For this, we perform experiments using four different values; 0.01, 0.001, 0.0001, and 0.00005. As evident from Fig. 7 below, 0.0001 performs best as compared with the others as it gives earlier convergence with high accuracy.

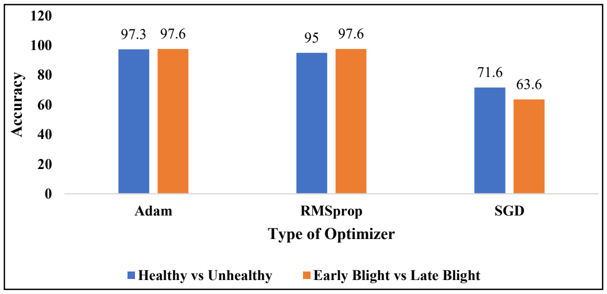


Fig. 6. Comparison of training performances using three different optimizers (Adam, RMSprop, and SGD) for both stage classifications

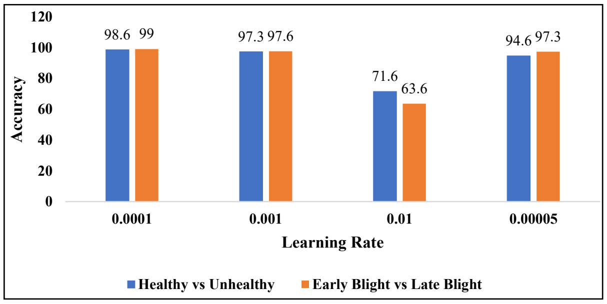


Fig. 7. Comparison of training performances using three different learning rates (0.0001, 0.001, 0.01, and 0.00005) for both stage classifications

Subsequently, to analyse the impact of network depth, we implement CNN architectures with 2, 3, and 4 convolutional layers. It was observed that the model with 3 convolutional layers yielded the most reliable and accurate results for both stage classifications. This provides the best trade-off between the complexity and performance, as large number of layers may capture more complex features but becomes computationally expensive. This analysis is provided in the Fig. 8 below.

To further justify the architecture design choices, we perform an ablation study by varying the number of layers and incorporating dropout regularization. The objective is to evaluate the impact of architectural depth and dropout on the model's performance and feature learning capability. Experiments are performed using one, two, and three layers, both with and without dropout, under identical training conditions. These results are summarized in Table 5 below. The ablation study is presented for healthy/ unhealthy classification only, as the other stage classification exhibits analogous trends and consistent performance behaviour.

It is evident from this table that increasing the number of layers improves accuracy up to a certain depth, beyond which the gain becomes marginal. Additionally, the

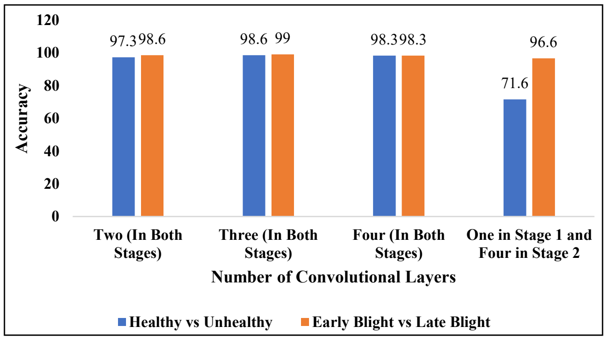


Fig. 8. Comparison of training performances using different convolutional layers for both stage classifications

Table 5. Ablation study results for varying model configurations (Values in Italics are for proposed model).

Layers	Dropout	Parameters	Accuracy
[32, 0, 0]	Yes	1,62,58,177	0.8100
[32, 0, 0]	No	1,62,58,177	0.9200
[32, 64, 0]	Yes	73,92,449	0.9567
[32, 64, 0]	No	73,92,449	0.9767
[16, 32, 64]	Yes	16,29,473	0.9567
[16, 32, 64]	No	16,29,473	0.9033
[32, 64, 128]	Yes	33,04,769	0.9767
[32, 64, 128]	No	33,04,769	0.9733

inclusion of dropout enhances generalization by reducing overfitting, particularly in deeper configurations. Additionally, the model with 3.3 million parameters is lightweight for deployment, trains fast (converge in 10 epochs), requires low memory and is capable for practical applications.

Thereafter, to qualitatively evaluate the model's predictions, we present sample outputs showing the actual and predicted labels for various leaf images. Fig. 9 below gives representative cases from each category: healthy, early blight, and late blight. From this figure, it is clear that the model reliably captures promising features, even in the instances where disease identification patterns are visually subtle.

Sample Image			
Actual Label	Healthy	Early Blight	Late Blight
Predicted Label	Healthy	Early Blight	Late Blight

Fig. 9. Sample images illustrating actual and predicted labels for three categories: Healthy, Early Blight, and Late Blight.

We also illustrate the figures for the training/ validation accuracy and loss curves for both classification stages. The accuracy and loss graphs for stage 1 are given in Fig. 10 and the ones for stage 2 are given in Fig. 11 below. These graphs confirm that the model is able to learn effectively over epochs without overfitting. From both the figures, we observe that the training and validation curves smoothly converge.

Finally, we plot the Receiver Operating Characteristic (ROC) curves for both

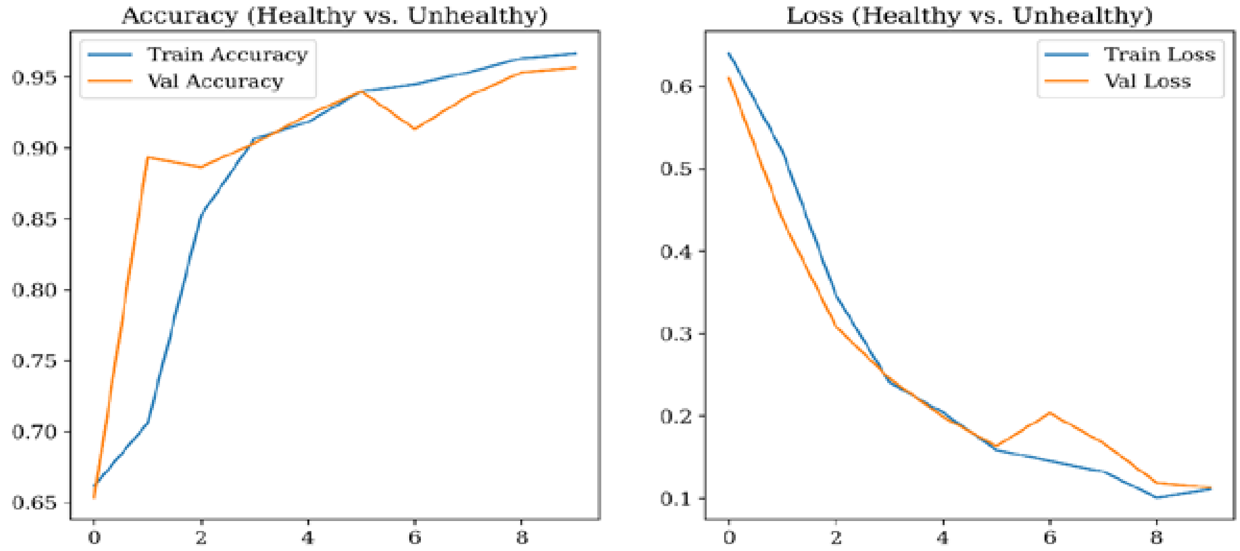


Fig. 10. Training and validation accuracy and loss curves for Stage 1 (Healthy vs. Unhealthy)

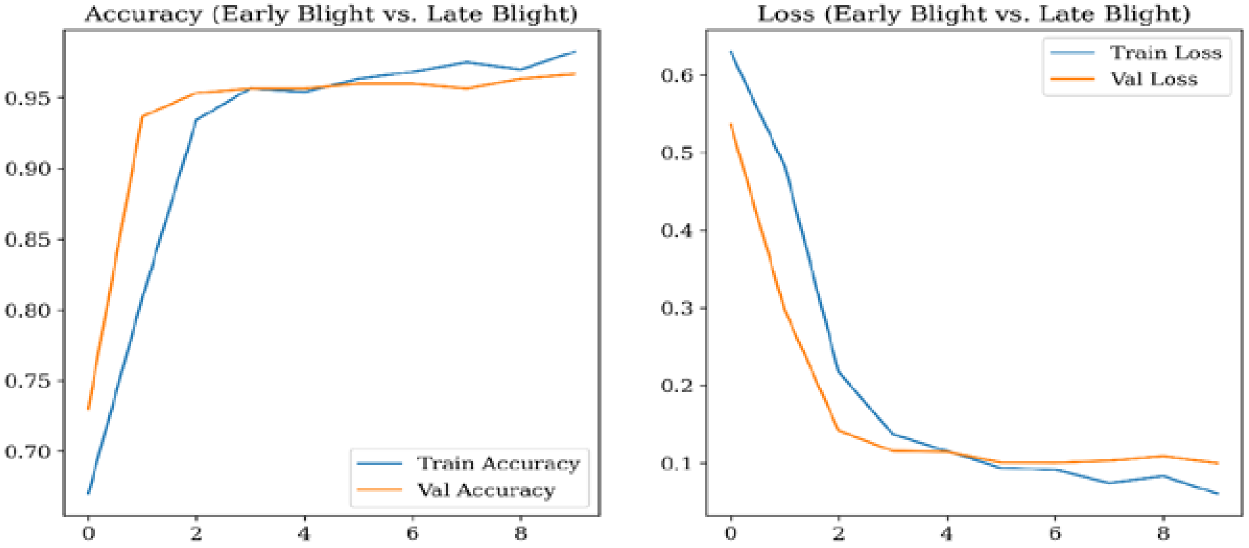
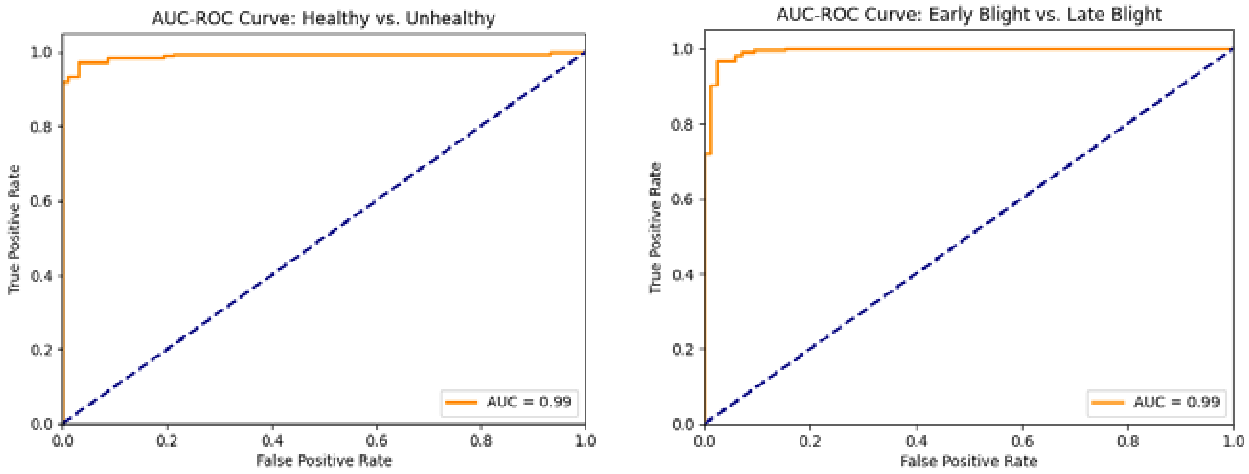


Fig. 11. Training/ Validation accuracy and loss curves for Stage 2 (Early Blight vs. Late Blight)

classification stages. And, the Area Under the Curve (AUC) was calculated to further monitor the discriminative ability of our proposed model. High AUC values in both cases indicate that the models perform well to distinguish between the respective classes. These analyses provide an additional layer of validation to our proposed model. These graphs are given in the Fig. 12 below.

In addition to the above results, a comparative evaluation is preformed against two state-of-the-art models, EfficientNet and DenseNet using the same dataset to further establish the robustness of the proposed approach. This evaluation is carried out for both the stages of proposed framework under identical experimental settings, including data preprocessing, augmentation, and training parameters. These results are given in the Table 6 below.



(a) Healthy vs. Unhealthy

(b) Early Blight vs. Late Blight

Fig. 12. Receiver Operating Characteristic (ROC) curves with corresponding Area Under the Curve (AUC) values for both Stage 1 (Healthy vs. Unhealthy) and Stage 2 (Early Blight vs. Late Blight) classifications.

Table 6. Performance comparison of the proposed model with EfficientNet and DenseNet across both stages.

Stage	Model	Accuracy	Precision	Recall	F1-Score
Healthy vs. Unhealthy	EfficientNet	0.6667	0.4445	0.6667	0.5333
	DenseNet	0.9667	0.9450	0.9555	0.9502
	Proposed	0.9833	0.9797	0.9700	0.9748
Early Blight vs. Late Blight	EfficientNet	0.5000	0.2500	0.5000	0.3334
	DenseNet	0.9611	0.9770	0.9444	0.9604
	Proposed	0.9900	0.9900	0.9900	0.9900

The results demonstrate that the proposed model consistently outperforms both EfficientNet and DenseNet in terms of accuracy and generalization, thereby highlighting its practical advantage and robustness even with a limited dataset.

Due to the limited size of the potato dataset used in this work, the proposed CNN model incorporates various strategies to ensure robustness. As mentioned earlier, we have applied three convolutional layers with increasing filter sizes, max pooling, and dropout layers to prevent overfitting. Additionally, in the preprocessing stage, the images were normalized to improve convergence stability. Furthermore, data augmentation including rotations, shifts, flips, zooming, and shearing is also applied to artificially increase dataset diversity. Learning curves show that training and validation accuracy increase steadily while maintaining a small gap, and loss decreases without divergence, indicating strong generalization beyond the limited training samples. These results demonstrate that the model is robust and capable of accurately classifying potato leaf diseases despite the small dataset size.

CONCLUSIONS AND FUTURE WORK

In this paper, we present a robust and efficient deep learning-based approach for automatic classification of potato leaf diseases using a two-stage CNN framework. The novelty of this work involves effective classification between healthy and unhealthy leaves in the first stage and early Blight and late blight in the second stage. Through extensive experimentation on 1,500 potato leaf images, we identify the optimal combination of model parameters; i.e., three convolutional layers, the Adam optimizer, and a learning rate of 0.0001 that give superior classification performance. Evaluation through confusion matrices, standard performance metrics, and ROC curves confirms the reliability and high accuracy of the model. We achieve around 98% accuracy in the first stage and 99% accuracy in the second stage of classification. The proposed system offers a scalable and accurate classification model for disease identification in agriculture, which can aid farmers and agronomists in identifying the diseases at an early stage.

While most of the recent studies focus on classification of leaf images, in future, we plan to localize the abnormality in the leaf using clustering techniques. This can provide more granular insights about the nature of the infection. We propose to explore the unsupervised clustering technique called spectral clustering that is known for its accurate results (Shastri A.A., *et al.*, 2019) Shastri A.A., Ahuja, Ratnaparkhe, & Busnel, 2021). Another future direction that can be explored involve implementation of multi-class classification systems which work for multiple crops types as well. This will make the model more generalize and efficient in real world settings. Finally, we can investigate the integration of CNNs with Internet of Things (IoT) devices, allowing real-time disease monitoring via mobile apps and smart agricultural sensors (Deng *et al.*, 2009).

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Department of Computer Science and Engineering for providing access to the "Next Generation Computing" laboratory, where all experimental work related to this research was carried out. The support and infrastructure offered by the department played a crucial role in the successful completion of this study.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ETHICAL APPROVAL

This article does not contain any studies with human participants or animals performed by any of the authors.

LITERATURE CITED

Adhikari, S. (2024). Advancements in Agricultural Technology: Vision Transformer-Based Potato Leaf Disease Classification. *Journal of Soft Computing Paradigm, 6*(2), 169-185.

Bajpai, A., Sahu, S., & Tiwari, N. K. (2025). Integrating Attention Mechanisms and Squeeze-and-Excitation Blocks for Accurate Potato Leaf Disease Detection. *Potato Research*, 1-21.

Chang, C. Y., & Lai, C. C. (2024). Potato leaf disease detection based on a lightweight deep learning model. *Machine Learning and Knowledge Extraction, 6*(4), 2321-2335.

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1251-1258). IEEE.

Chowdhury, N., Sultana, J., Rahman, T., Chowdhury, T., Khan, F. T., & Chakraborty, A. (2024). Potato leaf disease detection through ensemble average deep learning model and classifying the disease severity. *Indonesian Journal of Electrical Engineering and Computer Science, 35*(1), 494-502.

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). 1*, pp. 886--893. IEEE.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248--255). IEEE.

Fuadi, I., Putri, R. N., Nasien, D., & Oktarina, D. (2024). Deep Learning Approaches for Potato Leaf Disease Detection: Evaluating the Efficacy of Convolutional Neural Network Architectures. *Revue d'Intelligence Artificielle, 38*(2), 717.

Gülmez, B. (2024). A comprehensive review of convolutional neural networks based disease detection strategies in potato agriculture. *Potato Research*, 1-35.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70--90.

Lanjewar, M. G., Morajkar, P., & Payaswini, P. (2024). Modified transfer learning frameworks to identify potato leaf diseases. *Multimedia Tools and Applications, 83*(17), 50401-50423.

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science, 7*, 215232.

Muhammad, A. P. (2025, March 15). *Potato Leaf Disease Dataset*. Retrieved from Kaggle: https://www.kaggle.com/datasets/muhammadardiputra/potato-leaf-disease-dataset

Paul, H., Ghatak, S., Chakraborty, S., Pandey, S. K., Dey, L., Show, D., & Maity, S. (2024). A study and comparison of deep learning based potato leaf disease detection and classification techniques using explainable AI. *Multimedia Tools and Applications, 83*(14), 42485-42518.

Pujari, J. D., Yakkundimath, R., & Byadgi, A. S. (2015). Image processing based detection of fungal diseases in plants. *Procedia Computer Science, 46*, 1802--1808.

Radwan, M., Alhussan, A. A., Ibrahim, A., & Tawfeek, S. M. (2024). Potato Leaf Disease Classification Using Optimized Machine Learning Models and Feature Selection Techniques. *Potato Research*, 1-25.

Rashid, J., Khan, I., Ali, G., Almotiri, S. H., AlGhamdi, M. A., & Masood, K. (2021). Multi-level deep learning model for potato leaf disease recognition. *Electronics, 10*(17), 2064.

Raza, A., Pitafi, A. H., Shaikh, M. K., & Ahmed, K. (2025). Optimizing Potato Leaf Disease Recognition: Insights DENSE-NET-121 and Gaussian Elimination Filter Fusion. *Heliyon, 11*(3), e42318.

Reis, H. C., & Turk, V. (2024). Potato leaf disease detection with a novel deep learning model based on depthwise separable convolution and transformer networks. *Engineering Applications of Artificial Intelligence, 133*, 108307.

Shastri, A. A., Ahuja, K., Ratnaparkhe, M., & Busnel, Y. (2021). Probabilistically sampled and spectrally clustered plant species using phenotypic characteristics. *PeerJ, 9*, e11927.

Shastri, A. A., Ahuja, K., Ratnaparkhe, M., Shah, A., Gagrani, A., & Lal, A. (2019). Vector quantized spectral clustering applied to whole genome sequences of plants. *Evolutionary Bioinformatics, 15*, 1--7.

Shastri, A. A., Prajapati, Y., Katariya, H., Paliwal, M., & Sabale, K. (2025). Enhancing Clinical Outcomes Using Deep Learning Solution for Accurate Lung Cancer Classification. *Sensing and Imaging, 26*(1), 19.

Shastri, A. A., Tamrakar, D., & Ahuja, K. (2018). Density-wise two stage mammogram classification using texture exploiting descriptors. *Expert Systems with Applications, 99*, 71--82.

Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information processing in Agriculture, 4*(1), 41--49.

Sofuoğlu, C. İ., & Bırant, D. (2024). Potato plant leaf disease detection using deep learning method. *Journal of Agricultural Sciences, 30*(1), 153-165.

Strange, R. N., & Scott, P. R. (2005). Plant disease: A threat to global food security. *Anual Report of Phytopathology, 43*, 83-116.

Upadhyay, S., Jain, J., & Prasad, R. (2024). Early Blight and Late Blight Disease Detection in Potato Using Efficientnetb0. *International Journal of Experimental Research and Review, 38*, 15-25.

Wahabzada, M., Mahlein, A. K., Bauckhage, C., Steiner, U., Oerke, E. C., & Kersting, K. (2015). Plant disease detection from image data using deep learning. *AI and Society, 30*(3), 463--474.

Zhang, C., Wang, S., Wang, C., Wang, H., Du, Y., & Zong, Z. (2025). Research on a Potato Leaf Disease Diagnosis System Based on Deep Learning. *Agriculture, 15*(4), 424.

(MS Received : June 6, 2025; Accepted : November 10, 2025)