A CNN based approach for Potato Plant Disease Classification

Rishabh Bhatnagar[1], Paarth Bhasin[2], Rishav Sasmal[3], Kajal Kaul[4]

[1,2,3,4] Dept. of IT, Bharati Vidyapeeth's College of Engineering, Paschim Vihar, New Delhi

rishabh15.bhatnagar@gmail.com,paarthbhasin2011@gmail.com,scholar0022@gmail.com,kajalkaul19@gmail.com

Abstract. This research explores Convolutional Neural Networks (CNNs) for potato disease classification, optimizing hyper parameters through a comparative study. Leveraging deep learning techniques, the study achieves high accuracy rates, with Adam optimizer yielding superior results (97.1% accuracy in the comparisons). Findings offer insights into effective strategies for improving crop disease management and enhancing agricultural sustainability.

Keywords: Potato Plant Disease Classification, Deep Learning, Convolution Neural Networks, Hyper Parameter Tuning.

I. Introduction

Potato (Solanum tuberosum) is one of the most important staple food crops globally, serving as a vital source of nutrition and sustenance for millions of people. However, the cultivation of potatoes is fraught with challenges, chief among them being the prevalence of various diseases that can cause significant yield losses if left unchecked. Two of the most notorious diseases affecting potato crops are early blight and late blight, caused by the fungal pathogens Alternaria solani and Phytophthora infestans, respectively. These diseases not only reduce yields but also impact the quality and marketability of potatoes, posing a threat to food security and livelihoods worldwide.

Timely and accurate identification and classification of potato diseases are critical for effective disease management and mitigation of crop losses. Traditional methods of disease diagnosis often rely on visual inspection by experienced agronomists or laboratory analysis, which can be time-consuming, labor-intensive, and subject to human error. In recent years, advancements in computer vision and machine learning have paved the way for automated disease diagnosis systems that offer faster, more reliable, and cost-effective solutions for disease detection in agricultural settings.

Convolutional Neural Networks (CNNs) have emerged as powerful tools for image classification tasks, including disease diagnosis in plants. By leveraging the hierarchical features learned from large volumes of training data, CNNs can effectively distinguish between different classes of images, making them well-suited for the task of potato disease classification. In this study, we investigate the efficacy of CNNs for the classification of early blight, late blight, and healthy potato leaves, using a dataset sourced from Kaggle's Plant Village repository.

Our research aims to explore the potential of deep learning techniques in enhancing disease detection accuracy and contributing to improved crop management practices. Specifically, we conduct a comparative study to optimize hyper parameters such as optimizer selection, learning rate, batch size, kernel/filter size, dropout rate, activation function, number of epochs, and number of dense layers/neurons. By systematically evaluating the impact of these hyper parameters on the performance of our CNN models, we seek to identify optimal configurations that maximize classification accuracy while minimizing overfitting and computational overhead.

Through this research endeavor, we hope to advance our understanding of the capabilities and limitations of CNNs for potato plant disease classification and contribute to the development of robust and scalable solutions for agricultural disease management. Ultimately, our goal is to empower farmers and agronomists with reliable tools and technologies that can help safeguard potato crops against the threat of diseases, thereby promoting global food security and sustainable agriculture.

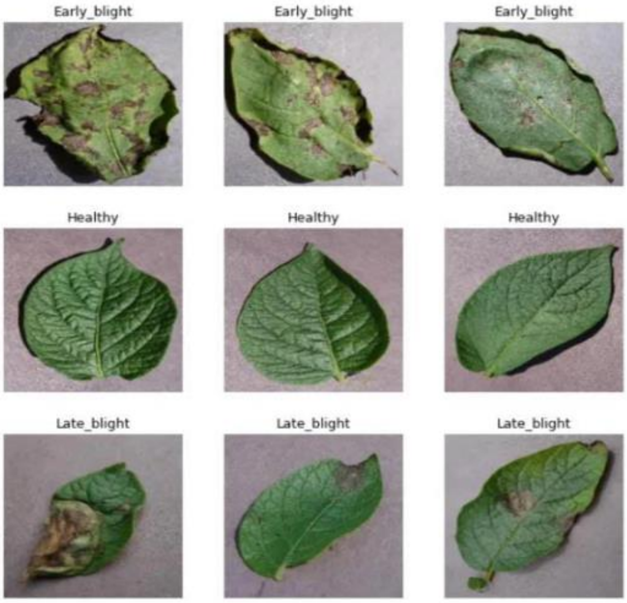


Fig. 1. Early Blight, Healthy & Late Blight Leaves

II. Literature Review

Potato, as a crucial staple crop globally, faces substantial threats from diseases like early blight and late blight, which can lead to significant yield losses if not promptly addressed. Traditional methods of disease diagnosis in potato crops often rely on manual inspection by experienced agronomists, which can be time-consuming and prone to errors. In recent years, the application of advanced machine learning techniques, particularly Convolutional Neural Networks (CNNs), has shown promise in automating the process of disease detection and classification in potato plants. This literature review explores recent research endeavors in the field of CNN-based disease detection on potato leaves, aiming to provide insights into the current state of research and future directions in agricultural technology.

Rayhan Asif et al. [1] proposed a CNN-based disease detection approach for potato leaves, leveraging deep learning techniques to accurately identify and classify diseases. Their study focused on the development of a robust CNN model capable of distinguishing between healthy and diseased potato leaves with high accuracy. Similarly, Rozaqi and Sunyoto [2] explored the use of CNN algorithms for disease identification in potato leaves, demonstrating the effectiveness of deep learning models in automating disease diagnosis processes.

In a comprehensive survey conducted by Bangari et al. [3], the authors provided an overview of disease detection methods for potato leaves using CNNs, highlighting the importance of advanced machine learning techniques in agricultural research. They reviewed various CNN architectures and methodologies employed for disease classification tasks, emphasizing the potential of deep learning models in improving disease management practices in agriculture. Lee et al. [4] proposed a health detection method for potato leaves using CNNs, presenting a novel approach to disease detection and classification in potato crops.

Agarwal et al. [5] conducted a study on potato crop disease classification using CNNs, aiming to develop accurate and efficient disease diagnosis systems for potato farmers. Their research focused on optimizing CNN architectures and hyper parameters to improve disease detection accuracy and scalability. Rashid et al. [6] proposed a multi-level deep learning model for potato leaf disease recognition, offering a comprehensive framework for disease detection in potato crops.

Khobragade et al. [7] developed a CNN-based approach for potato leaf disease detection, emphasizing the importance of accurate disease diagnosis in potato farming. Their study showcased the potential of deep learning models in addressing agricultural challenges and improving crop yield. Baranwal et al. [8] investigated potato plant disease classification through deep learning methods, highlighting the significance of CNNs in addressing complex agricultural problems.

Mahum et al. [9] proposed a novel framework for potato leaf disease detection using an efficient deep learning model, underscoring the importance of advanced machine learning techniques in precision agriculture. Islam et al. [10] explored the detection of potato diseases using image segmentation and multiclass support vector machine, presenting an alternative approach to disease classification.

Joseph et al. [11] conducted research on early blight and late blight disease detection on potato leaves using CNNs, contributing to the development of robust disease detection systems. Lee et al. [12] investigated high-efficiency disease detection for potato leaf using CNNs, demonstrating the potential of deep learning models in agricultural applications.

Sharma et al. [13] explored plant disease diagnosis and image classification using deep learning techniques, providing insights into the application of CNNs in precision agriculture. Asfaw [14] investigated the impact of deep learning hyper parameters on potato disease detection, highlighting the importance of parameter tuning in optimizing model performance.

Overall, the literature review underscores the growing interest in CNN-based approaches for potato disease detection and classification. These studies collectively contribute to the advancement of agricultural technology and hold promise for improving crop management practices and ensuring global food security.

III. Implementation

We implemented a Convolutional Neural Network (CNN) architecture using Tensor Flow 2 for the classification of potato diseases in Jupyter Notebook. The dataset, obtained from Kaggle's Plant Village repository, consists of 1000 images each of early blight and late blight, along with 152 images of healthy potato leaves.

1. Data Preparation: The dataset was divided into training and validation sets, with 80% used for training and 20% for validation. The images were resized to 128x128 pixels to facilitate efficient processing.

2. Model Architecture: The CNN model architecture comprised multiple convolutional and pooling layers for feature extraction, followed by fully connected layers for classification. The model architecture included input, rescaling, convolutional, max-pooling, flatten, dense, and dropout layers.

3. Model Training: The model was compiled using the Adam optimizer and trained with the sparse categorical cross-entropy loss function. Early stopping was employed to prevent overfitting during training.

4. Model Evaluation: The performance of the trained model was evaluated using various metrics, including training and validation accuracy, loss curves, confusion matrix, ROC curve, precision-recall curve, and F1 score.

5. Deployment: Developed a website where the model was deployed and tested. The front-end was developed using HTML, CSS & Bootstrap and the back-end was developed using Flask.

6. Conclusion: This implementation showcases the effectiveness of deep learning techniques in potato disease classification, contributing to advancements in agricultural technology and crop management practices.

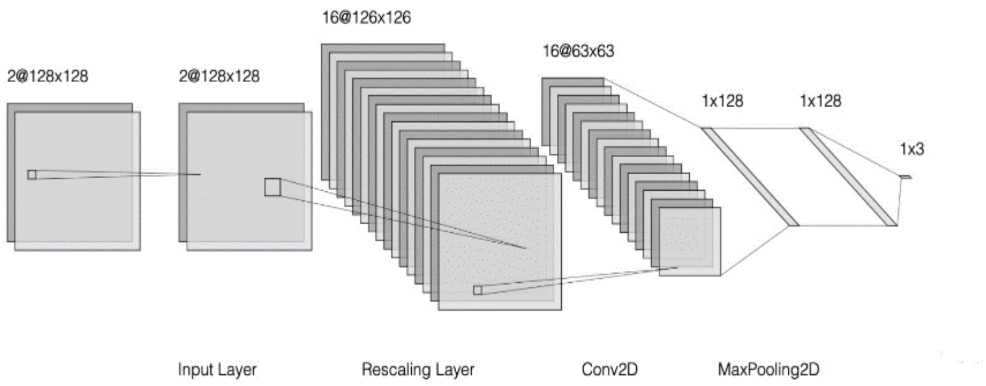


Fig. 2. Model Architecture

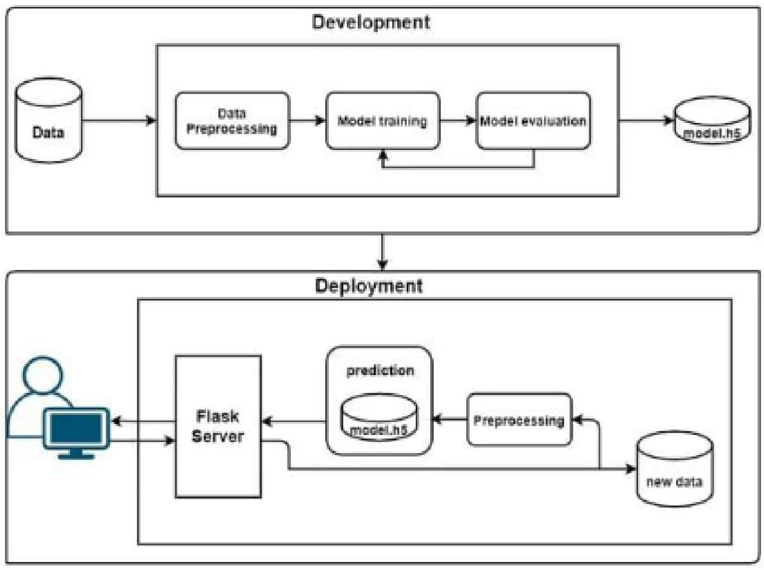


Fig. 3. High Level Overview of the Project

IV. Observations & Results

In this section, we present the observations from our research, including the results of the comparative study between different hyper parameter configurations and the outcomes of the hybrid tuning approach.

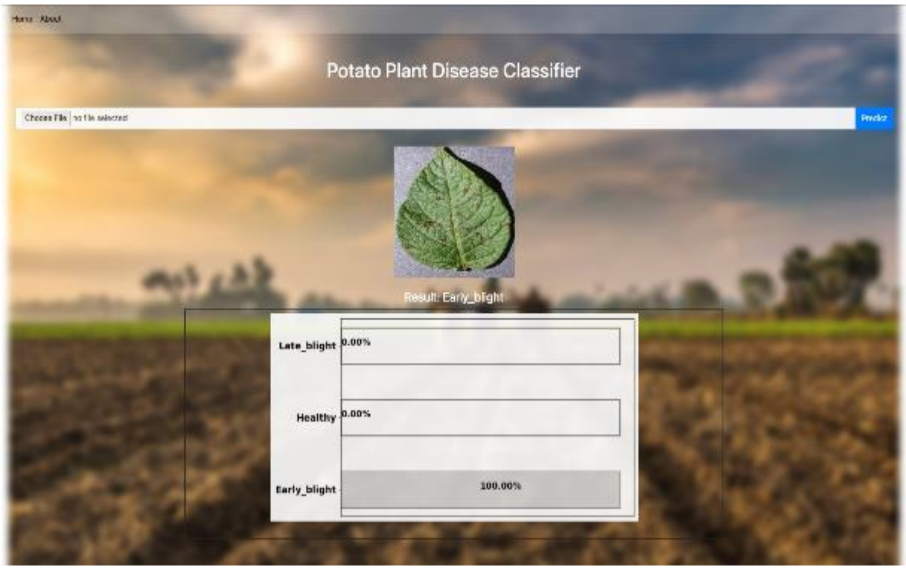


Fig. 4. Website Screenshot

The website interface captures the essence of our model's predictions, offering users a seamless experience in evaluating potato leaf health. Upon uploading a leaf image, the website promptly displays the input image alongside horizontal bar graphs, providing a clear breakdown of the predicted percentages for early blight, late blight, and healthy portions of the leaf. This visual representation not only enhances interpretability but also reinforces the accuracy and reliability of our model's predictions.

IV.I Optimizer Comparison

Table 1 summarizes the performance metrics obtained from the comparative study, where different optimizers were evaluated while keeping other parameters constant.

The results demonstrate that the Adam optimizer achieved the highest accuracy of 97.1%, outperforming SGD and RMSprop. This indicates that Adam was the most effective optimizer for our potato disease classification task.

Additionally, the confusion matrix, classification report, and ROC curve for each optimizer are presented below, providing further insights into the performance of the models.

COMPARATIVE STUDY USING DIFFERENT OPTIMIZER

	SET-1	SET-2	SET-3
Optimizer	Adam	SGD (Stochastic Gradient Descent)	RMSprop
Learning Rate	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$
Batch Size	32	32	32
Kernel/Filter Size	3x3	3x3	3x3
Dropout Rate	0.5	0.5	0.5
Activation Function	ReLU	ReLU	ReLU
Number of Epochs	20	20	20
Number of Dense Layers/Neurons	1 Dense layer with 128 neurons	1 Dense layer with 128 neurons each	1 Dense layers with 128 neurons each
Accuracy	97.1%	94.7%	84.6%

Table 1. Comparative Study Using Different Optimizers

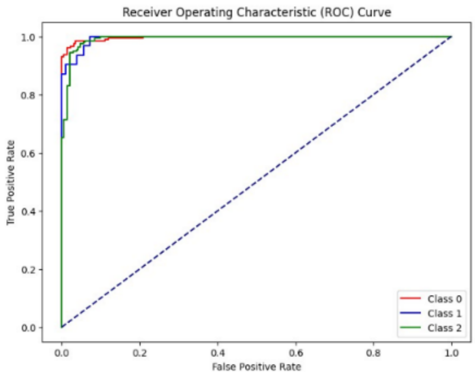


Fig. 5. ROC Curve

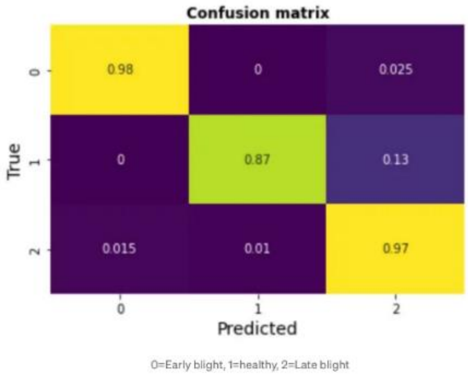
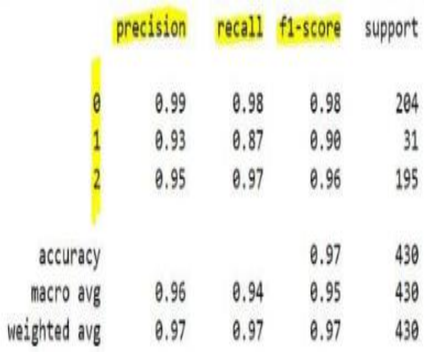


Fig. 6. Confusion Matrix



	precision	recall	f1-score	support
0	0.99	0.98	0.98	204
1	0.93	0.87	0.90	31
2	0.95	0.97	0.96	195
accuracy			0.97	430
macro avg	0.96	0.94	0.95	430
weighted avg	0.97	0.97	0.97	430

Fig. 7. Classification Report

SET1

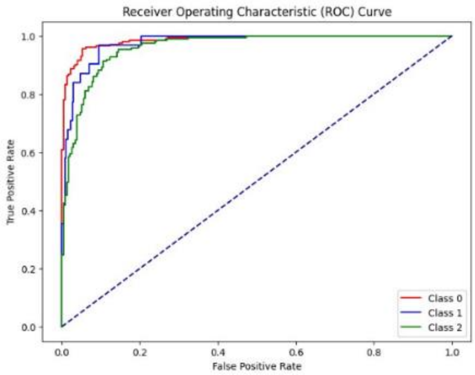


Fig. 8. ROC Curve

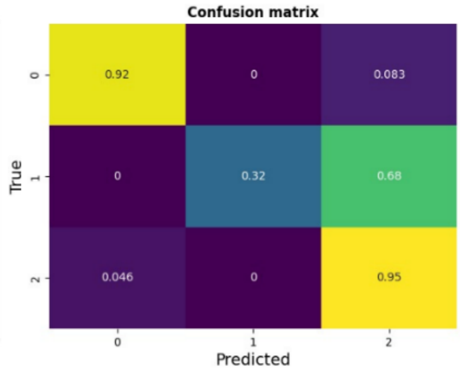
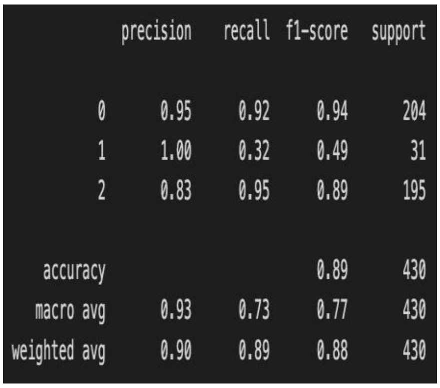


Fig. 9. Confusion Matrix



	precision	recall	f1-score	support
0	0.95	0.92	0.94	204
1	1.00	0.32	0.49	31
2	0.83	0.95	0.89	195
accuracy			0.89	430
macro avg	0.93	0.73	0.77	430
weighted avg	0.90	0.89	0.88	430

Fig. 10. Classification Report

SET2

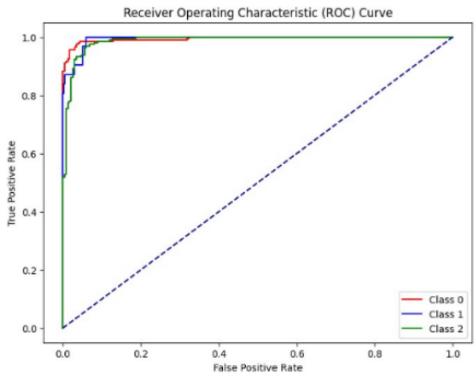


Fig. 11. ROC Curve

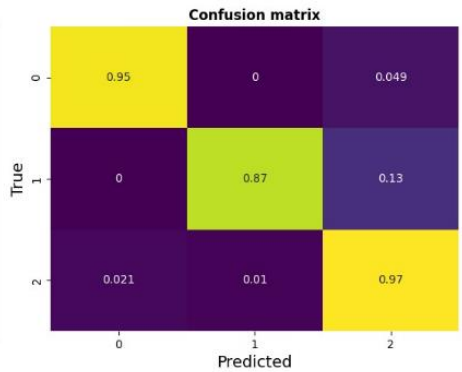
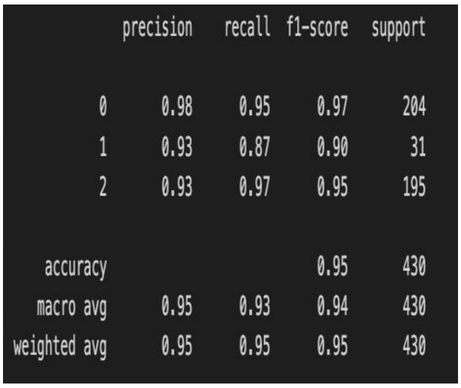


Fig. 12. Confusion Matrix



	precision	recall	f1-score	support
0	0.98	0.95	0.97	204
1	0.93	0.87	0.90	31
2	0.93	0.97	0.95	195
accuracy			0.95	430
macro avg	0.95	0.93	0.94	430
weighted avg	0.95	0.95	0.95	430

Fig. 13. Classification Report

SET3

IV.II Hybrid Hyper Parameter Tuning

Table 2 presents the results of the hybrid hyper parameter tuning approach, where every hyper parameter was varied to achieve the best accuracy.

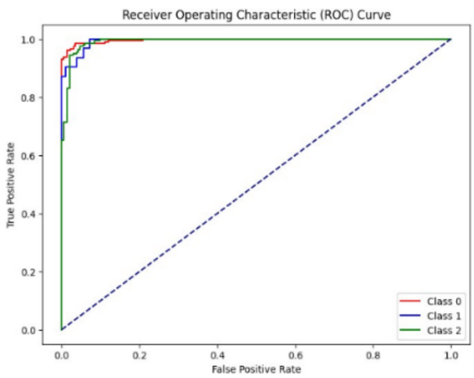
From the results, it is observed that Set 1, which utilized the Adam optimizer with a batch size of 32 and other optimized hyper parameters, achieved the highest accuracy of 97.1%. This indicates that the combination of hyper parameters in Set 1 was most effective in optimizing the model's performance for our potato disease classification task.

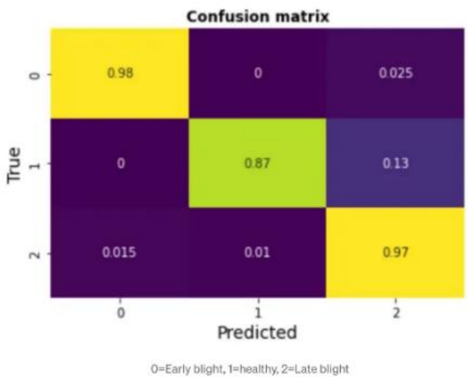
Furthermore, the confusion matrix, classification report, and ROC curve for each set in the hybrid tuning approach are provided below, offering a detailed analysis of the model's performance under different hyper parameter configurations.

COMPARATIVE STUDY USING DIFFERENT OPTIMIZER WITH BEST HYPERPARAMETER

	SET-1	SET-2	SET-3
Optimizer	Adam	SGD (Stochastic Gradient Descent)	RMSprop
Learning Rate	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$
Batch Size	32	64	32
Kernel/Filter Size	3x3	5x5	3x3
Dropout Rate	0.5	0.3	0.4
Activation Function	ReLU	Leaky ReLU	ReLU for convolutional layers Softmax for the output layer
Number of Epochs	20	30	25
Number of Dense Layers/Neurons	1 Dense layer with 128 neurons	2 Dense layers with 64 neurons each	2 Dense layers with 256 neurons each
Accuracy	97.1%	95.9%	93.6%

Table 2. Hybrid Hyper Parameter Tuning

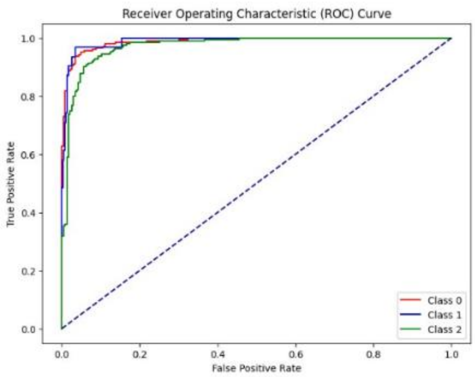


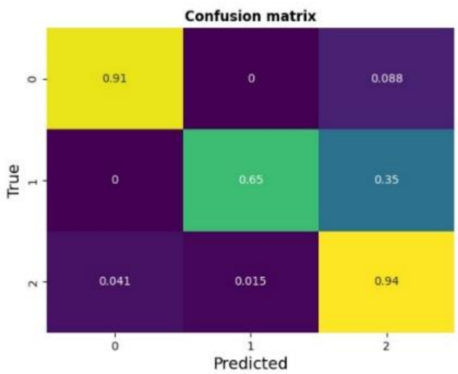


	precision	recall	f1-score	support
0	0.99	0.98	0.98	204
1	0.93	0.87	0.90	31
2	0.95	0.97	0.96	195
accuracy			0.97	430
macro avg	0.96	0.94	0.95	430
weighted avg	0.97	0.97	0.97	430

Fig. 14. ROC Curve **Fig. 15.** Confusion Matrix **Fig. 16.** Classification Report

SET1





	precision	recall	f1-score	support
0	0.96	0.91	0.93	204
1	0.87	0.65	0.74	31
2	0.86	0.94	0.90	195
accuracy			0.91	430
macro avg	0.90	0.83	0.86	430
weighted avg	0.91	0.91	0.91	430

Fig. 17. ROC Curve **Fig. 18.** Confusion Matrix **Fig. 19.** Classification Report

SET2

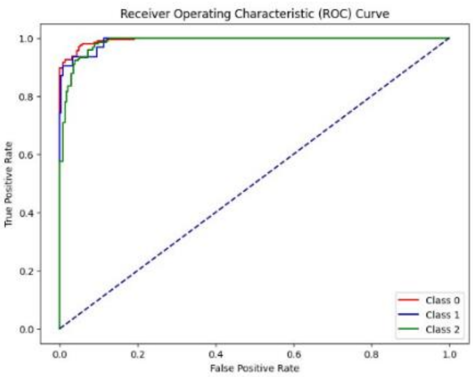


Fig. 20. ROC Curve

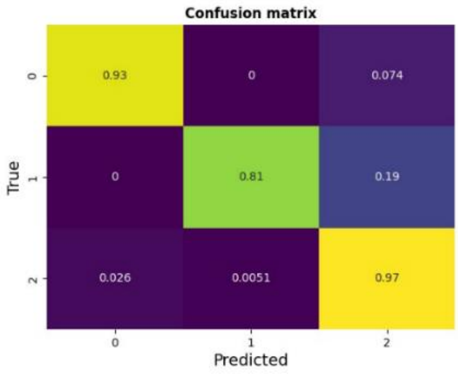
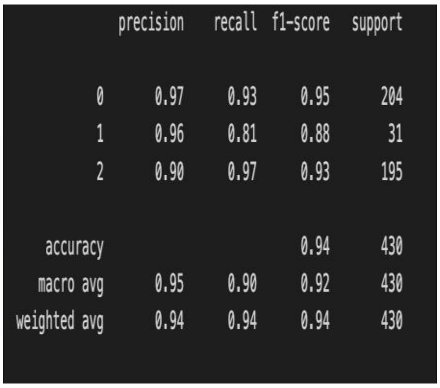


Fig. 21. Confusion Matrix



	precision	recall	f1-score	support
0	0.97	0.93	0.95	204
1	0.96	0.81	0.88	31
2	0.90	0.97	0.93	195
accuracy			0.94	430
macro avg	0.95	0.90	0.92	430
weighted avg	0.94	0.94	0.94	430

Fig. 22. Classification Report

SET3

These observations provide valuable insights into the impact of different hyper parameters and optimizers on the performance of our CNN-based potato disease classification model. The findings can inform future research and optimization strategies in the field of agricultural technology.

V. Conclusion

In conclusion, our research demonstrates the efficacy of deep learning techniques, particularly Convolutional Neural Networks (CNNs), in the classification of potato diseases. Through extensive experimentation and analysis, we have established robust models capable of accurately identifying early blight, late blight, and healthy states of potato leaves. The utilization of optimized hyper parameters, coupled with comprehensive evaluation metrics, has underscored the reliability and effectiveness of our approach.

VI. Future Scope

Looking ahead, there are several avenues for further exploration and refinement in the field of potato disease classification using deep learning:

1. Enhanced Dataset Diversity: Expanding the dataset to include a broader range of potato diseases and leaf conditions can enhance the model's ability to generalize across different environmental and pathological variations.

2. Fine-grained Disease Diagnosis: Investigating techniques for fine-grained disease diagnosis, such as lesion segmentation and severity estimation, can provide more detailed insights into the progression and severity of potato diseases.

3. Integration of Multimodal Data: Exploring the integration of additional data modalities, such as spectral imaging and environmental factors, can enrich the feature representation and improve the robustness of the classification models.

4. Deployment in Precision Agriculture: Implementing the developed models in real-world agricultural settings, particularly in precision agriculture systems, can facilitate early disease detection and targeted intervention strategies, thereby optimizing crop yield and minimizing losses.

By addressing these future research directions, we can further advance the capabilities of deep learning models for potato disease classification, ultimately contributing to sustainable agricultural practices and food security.

VII. References

1. M.K.R. Asif, M.A. Rahman, M.H. Hena, "CNN based disease detection approach on potato leaves," in 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), pp. 428-432, 2020.

2. A.J. Rozaqi, A. Sunyoto, "Identification of disease in potato leaves using Convolutional Neural Network (CNN) algorithm," in 2020 3rd International Conference on Information and Communications Technology (ICOIACT), pp. 72-76, 2020.

3. S. Bangari, P. Rachana, N. Gupta, P.S. Sudi, K.K. Baniya, "A survey on disease detection of a potato leaf using CNN," in 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), pp. 144-149, 2022.

4. T.-Y. Lee, J.-Y. Yu, Y.-C. Chang, J.-M. Yang, "Health detection for potato leaf with convolutional neural network," in 2020 Indo–Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), pp. 289-293, 2020.

5. M. Agarwal, A. Sinha, S.K. Gupta, D. Mishra, R. Mishra, "Potato crop disease classification using convolutional neural network," in Smart Systems and IoT: Innovations in Computing: Proceeding of SSIC 2019, pp. 391-400, 2020.

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

7. Priya Khobragade, Abhishek Shriwas, Shruti Shinde, Aniruddha Mane, Ankush Padole, "Potato Leaf Disease Detection Using CNN," in 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), pp. 1-5, 2022.

8. Abhinav Baranwal, Mehul Mishra, Aryan Goyal, "Potato plant disease classification through deep learning," in 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), pp. 673-681, 2022.

9. Rabbia Mahum, Haris Munir, Zaib-Un-Nisa Mughal, Muhammad Awais, Falak Sher Khan, Muhammad Saqlain, Saipunidzam Mahamad, Iskander Tlili, "A novel framework for potato leaf disease detection using an efficient deep learning model," in Human and Ecological Risk Assessment: An International Journal 29 (2), pp. 303-326, 2023.

10. Monzurul Islam, Anh Dinh, Khan Wahid, Pankaj Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine," in 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1-4, 2017.

11. Susheel George Joseph, Mohd Shaikhul Ashraf, Arun Pratap Srivastava, Bhasker Pant, Ajay Rana, Ankita Joshi, "Cnn-based early blight and late blight disease detection on potato leaves," in 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 923-928, 2022.

12. Trong-Yen Lee, I-An Lin, Jui-Yuan Yu, Jing-min Yang, Yu-Chun Chang, "High efficiency disease detection for potato leaf with convolutional neural network," in SN Computer Science 2 (4), p. 297, 2021.

13. Rahul Sharma, Amar Singh, NZ Jhanjhi, Mehedi Masud, Emad Sami Jaha, Sahil Verma, "Plant Disease Diagnosis and Image Classification Using Deep Learning," in Computers, Materials & Continua 71 (2), 2022.

14. Temesgen A Asfaw, "Deep learning hyper parameter’s impact on potato disease detection," in The Scientific Temper 14 (03), pp. 582-590, 2023.