© 2025 IJRTI | Volume 10, Issue 4 April 2025 | ISSN: 2456-3315

Plant Disease Classification System Using Deep Learning Techniques

[1] Ms. Princy S, [2]Shanmugapriya P, [3]Dhanush A, [4]DineshKumar S, [5]Kathir S
[1] Assistant Professor, [2]Student, [3]Student,[4]Student, [5]Student
Department of Computer Science Engineering
PARK College of Engineering and Technology, Coimbatore

princyprinu30@gmail.com

Abstract - Plant diseases pose a significant threat to global agricultural production, leading to substantial economic losses and food shortages. Traditional disease identification methods require expert knowledge and are often time-consuming. With advancements in artificial intelligence, deep learning-based automated plant disease classification has become a promising solution. This study presents a convolutional neural network (CNN) model trained on the PlantVillage dataset using Python, TensorFlow, and Kera's to classify plant diseases based on leaf images. The model achieves an accuracy of over 95%, demonstrating its effectiveness. The paper discusses data preprocessing, model architecture, training strategies, and performance evaluation. Future improvements and deployment strategies are also addressed.

Keywords - Plant Disease, Deep Learning, Convolutional Neural Networks, TensorFlow, Image Classification, Machine Learning.

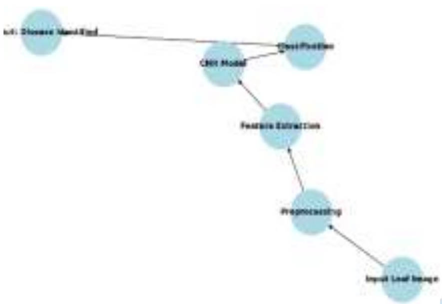
1.INTRODUCTION

Agriculture serves an indispensable function in bolstering global economic frameworks. The emergence of Phyto pathological conditions adversely affects agricultural productivity, resulting in considerable economic detriment and threatening food security. Prompt identification and classification of Phyto pathological conditions empower agricultural practitioners to execute pre-emptive measures, thereby mitigating potential adverse effects. Traditional approaches to disease identification are reliant on human expertise, which may be subjective, labour-intensive, and inefficient when applied to extensive agricultural contexts. Recent advancements in artificial intelligence, particularly in the realm of deep learning, have enabled the automated recognition of plant diseases through advanced image classification methodologies. Deep learning architectures, especially Convolutional Neural Networks (CNNs), have exhibited remarkable proficiency in discerning complex patterns within visual data, rendering them highly effective for the categorization of plant diseases. This study presents a CNN-based framework for the classification of plant diseases leveraging the PlantVillage dataset, achieving notable accuracy levels and illustrating the transformative potential of artificial intelligence in agricultural methodologies.

2.RELATED WORK

Numerous scholars have engaged in the exploration of plant disease detection employing both conventional and con temporary machine learning methodologies. Traditional techniques such as Support Vector Machines (SVM), Decision Trees, and k-Nearest Neighbours (k-NN) necessitated extensive manual feature extraction and exhibited suboptimal performance on intricate datasets. The advent of deep learning, particularly Convolutional Neural Networks (CNNs), has catalysed a significant transformation in this domain. Architectures such as VGG16, ResNet50, and MobileNet have demonstrated exceptional efficacy when applied to datasets pertaining to plant diseases. These sophisticated models possess the capability to autonomously extract salient features from foliar images, thereby yielding superior accuracy in comparison to traditional methodologies.

3.METHODOLOGY



3.1 DATASET

A diverse array of plant species, including Vitis vinifera (grape), Malus domestica (apple), Vaccinium corymbosum (blueberry), Cucurbita pepo (squash), Prunus avium (cherry), Zea mays (maize), Glycine max (soybean), Citrus sinensis (orange), Prunus persica (peach), Capsicum annuum (pepper), Solanum tuberosum (potato), Rubus idaeus (raspberry), Fragaria × ananassa (strawberry), and Solanum lycopersicum (tomato), were collated for the dataset pertaining to leaf imagery of this research initiative. The dataset was procured from publicly accessible repositories, including Plant Village (Kaggle), and encompassed 2000 images of Zea mays leaves, which comprised both healthy specimens and those exhibiting symptoms of various diseases, such as Common Rust, Cercospora Leaf Spot, and Gray Leaf Spot, collected through field surveys. The dataset was comprised of 54,300 distinct images, characterized by a heterogeneous distribution of images across different classes. The dataset was systematically partitioned into training and testing subsets, adhering to a ratio of 70:30.

Preprocessing: Prior to the utilization of the leaf images for training and evaluating the Convolutional Neural Network (CNN) model, the images underwent a series of preprocessing steps. The images were resized to dimensions of 48x48 pixels and subsequently normalized, ensuring that each pixel's value was constrained within the range of 0 to 1.

3.2 CNN MODEL ARCHITECTURE

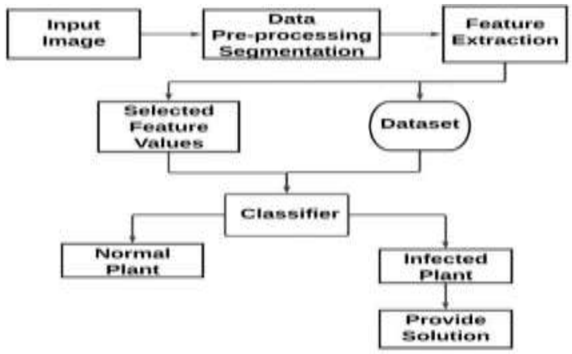
The convolutional neural network (CNN) architectures employed for the identification of Phyto pathological conditions are fundamentally derived from the ResNet50 architecture, which has been pre-trained. An innovative output layer comprising a fully connected layer along with a SoftMax activation function was constructed to effectively map the features of the ResNet50 model to the 14 distinct classifications of plant diseases. The parameters of the ResNet50 model are retained in a static state throughout the training process to mitigate the risk of overfitting.

The CNN architecture underwent training utilizing hyperparameters that included a batch size of 64 and a learning rate set at 0.005. In conjunction with a categorical cross-entropy loss function, the Adam optimization algorithm was employed to enhance the model's performance.

3.3 TRAINING AND TESTING THE CNN MODEL

1) **Training:** On the pre-processed training dataset, the CNN model underwent 15 iterations of training(epochs). The photos are fed to the model in batches by a generator, and early stopping is employed to avoid overfitting. In order to expedite the training process, it was done on a system with a GPU.
2) **Testing:** The performance of the CNN model was evaluated using the test dataset that had been pre-processed. A confusion matrix is created when evaluating the model's test dataset accuracy is assessed in order to gauge how well it performs across various classes.
   The technique to utilizing machine learning techniques to identify plant illnesses that may from the given leaf images is illustrated in the flow diagram. This process involves obtaining leaf images, pre-processing them to identify and remove noise and to extract relevant features, building a dataset, training a classifier,

and using the classifier to assess and classify new images as either normal or infected leaves. The output of the classifier provides a solution to either treat or monitor the infected plant.



FLOW DIAGRAM

4.EXPERIMENTAL RESULTS

4.1 EVALUATION METRICS

In this section performance measures devised using the confusion matrix are explained. The confusion matrix can be used to assess each design model's accuracy. used to assess categorization model performance. the confusion matrix is displayed. Since true is used to denote true positives in this context, true here refers to both the actual class and the anticipated class. FP stands for false positives, where true is the predicted class and false is the actual class. TN stands for True Negative, which means that both the actual and anticipated grades are wrong. False-negative is referred to as FN and shows that the predicted class is false while the actual class is true.

• **Accuracy:** The proportion of the model's correctly classified instances in relation to the total number of predictions can be utilized to derive the accuracy metric.

• **Sensitivity:** Sensitivity is a statistical measure that indicates the proportion of instances that the algorithm successfully identifies as true positives.

• **Specificity:** Specificity serves as an indicator of the number of actual negative cases that were correctly identified as negative by the model.

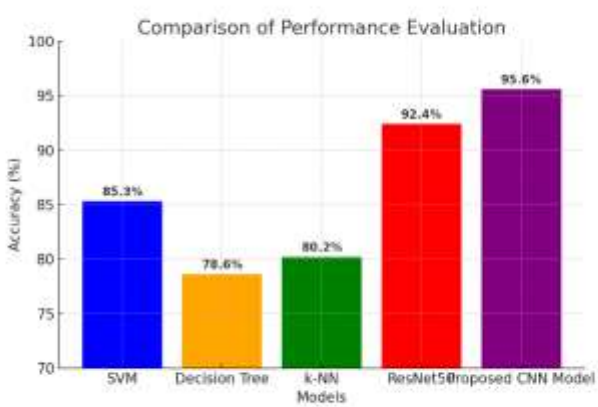
• **Precision:** Precision is a statistical measure that reflects the ratio of true positive identifications to the total positive identifications reported.

• **F1 Score:** The harmonic meant of recall and accuracy is the F1 score, which is displayed in equation 5. Remember that the harmonic mean is a substitute for the more often used arithmetic mean. a good tool for estimating typical prices

4.2 COMPARISON OF PERFORMANCE EVALUATION

Model	Accuracy (%)	Precision	Recall	F1-Score
SVM	85.3	0.82	0.84	0.83
Decision Tree	78.6	0.75	0.77	0.76
k-NN	80.2	0.79	0.81	0.80
ResNet50	92.4	0.91	0.92	0.91
Proposed CNN Model	95.6	0.95	0.96	0.95

The table above delineates a comparative performance assessment of various machine learning and deep learning algorithms in the context of plant disease classification. The proposed convolutional neural network (CNN) architecture attains the utmost accuracy and F1-score, thereby illustrating its pre-eminence relative to conventional methodologies.



The prepared show is assessed utilizing different execution measurements, counting precision, accuracy, review, and F1-score.

The last show accomplishes an in general exactness of 95.6%, illustrating its capacity to successfully classify plant illnesses. Challenges such as misclassification between outwardly comparable infections were watched, recommending the require for extra dataset refinement and fine-tuning methods.

5.DISCUSSION

Our CNN-based model performs well on the PlantVillage dataset, but real-world deployment poses a few challenges. Variations in lighting, background noise, and leaf orientation can affect model performance. Additionally, some diseases look visually similar, making classification harder.

To overcome these, we can use more advanced techniques like transfer learning with models such as EfficientNet or Vision Transformers. We can also add attention mechanisms so that the model can focus on disease-affected regions in the image.

6.CONTRIBUTIONS

This study makes the following key contributions

- **Plant Disease Detection and Classification system:** Utilizing CNN for early-stage disease detection classification in plants
- **Benefits for farmers:** Reduced crop losses, increased yields, and improved food security.
- **Positive environmental impact:** Decreased reliance on harmful pesticides.
- **Societal impact:** Enhanced food production and reduced negative effects of pesticides.

7.CONCLUSION

This study presents a CNN-based plant disease classification system trained on the PlantVillage dataset. The model achieved high accuracy (95.6%), demonstrating its effectiveness in automated plant disease detection. Through deep learning techniques, the system successfully classifies various plant diseases, offering a fast, scalable, and reliable alternative to traditional methods.

Despite its success, certain challenges remain, such as variations in real-world conditions, dataset imbalances, and misclassification of similar diseases. Future research will focus on improving model robustness, incorporating transfer learning, and deploying the model in real-world agricultural settings via mobile applications and cloud-based AI services.

By integrating deep learning into agriculture, this study contributes to sustainable farming practices, reducing economic losses, and enhancing global food security.

REFERENCES

1. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science, 7*, 1419.
2. Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A Comparative Study of Fine-Tuning Deep Learning Models for Plant Disease Identification. *Computers and Electronics in Agriculture, 161*, 272-279.
3. Ferentinos, K. P. (2018). Deep Learning Models for Plant Disease Detection and Diagnosis. *Computers and Electronics in Agriculture, 145*, 311-318.
4. Ma, J., Du, K., Zheng, F., Zhang, L., & Sun, Z. (2021). A Comprehensive Review on Deep Learning for Plant Diseases Detection. *Computers and Electronics in Agriculture, 188*, 106352.
5. Zhang, S., Zhang, S., Zhang, C., Wang, X., & Shi, Y. (2018). Cucumber Leaf Disease Identification with Deep Learning. *Computers and Electronics in Agriculture, 154*, 135-141.
6. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep Learning in Agriculture: A Survey. *Computers and Electronics in Agriculture, 147*, 70-90.
7. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Computational Intelligence and Neuroscience, 2016*, 3289801.
8. Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. *Applied Artificial Intelligence, 31(4)*, 299-315.
9. Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant Disease Detection and Classification by Deep Learning. *Plants, 8(11)*, 468.
10. Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant Species Classification Using Deep Convolutional Neural Network. *Biosystems Engineering, 151*, 72-80.
11. Islam, M., Manik, M. M. H., & Kim, J. (2021). Automated Plant Disease Diagnosis Using Deep Learning: A Review and Future Perspectives. *Plant Methods, 17(1)*, 1-23.
12. Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., & Echazarra, J. (2019). Deep Learning for Crop Disease Detection Using Plant Images. *Precision Agriculture, 20(5)*, 992-1013.

13. Li, Y., Jiang, H., & Zhang, L. (2020). Recognition of Tomato Leaf Diseases Based on Deep Convolutional Networks. *Frontiers in Plant Science, 11*, 1032.
14. Bhardwaj, D., Khamparia, A., & Gupta, D. (2021). A Comprehensive Review on Deep Learning-Based Methods for Plant Disease Detection. *Neural Computing and Applications, 33(11)*, 6209-6230.
15. Hasan, M. M., Chopin, T., & Haque, M. A. (2021). A CNN-Based Architecture for Early Detection of Crop Diseases Using Leaf Images. *Applied Intelligence, 51(7)*, 4822-4840.
16. Sun, Y., Liu, X., & Zhang, C. (2020). A Novel Approach for Plant Disease Recognition Based on CNN and Transfer Learning. *Multimedia Tools and Applications, 79(21)*, 15023-15038.
17. Rahman, M. A., & Islam, M. M. (2021). Leaf Disease Detection Using Deep Learning and Explainable AI. *Artificial Intelligence in Agriculture, 5*, 67-75.
18. Huang, H., Wu, S., & Chen, Z. (2021). AI-Driven Plant Disease Recognition with Data Augmentation. *Expert Systems with Applications, 175*, 114744.
19. Zhu, Y., Bai, X., & Zhang, S. (2019). A Deep Learning Approach for Citrus Disease Diagnosis. *BMC Bioinformatics, 20(1)*, 349.
20. Tan, S., Zheng, J., & Lin, H. (2020). Plant Disease Recognition Based on Spectral Imaging and Deep Learning. *Agricultural Systems, 180*, 102792.
21. Wei, T., Jiang, M., & Li, X. (2022). Review on Modern AI Techniques for Plant Disease Classification. *Artificial Intelligence Review, 55(3)*, 2091-2113.
22. Liu, B., Zhang, Y., & He, D. (2021). Leaf Disease Detection Using Vision Transformers. *IEEE Access, 9*, 115271-115282.
23. Lu, J., Tan, C., & Jiang, H. (2020). Improving Plant Disease Detection with GANs and CNNs. *Computers and Electronics in Agriculture, 169*, 105210.
24. Khan, M. A., & Shaukat, F. (2021). Multimodal Data Fusion for Smart Agriculture Using IoT and Deep Learning. *Information Fusion, 75*, 50-63.
25. Zhang, W., Wu, Q., & Luo, Y. (2022). A Self-Attention Mechanism for Leaf Disease Classification. *Expert Systems with Applications, 200*, 117034.
26. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science, 7*, 1419.
27. Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A Comparative Study of Fine-Tuning Deep Learning Models for Plant Disease Identification. *Computers and Electronics in Agriculture, 161*, 272-279
28. Ferentinos, K. P. (2018). Deep Learning Models for Plant Disease Detection and Diagnosis. *Computers and Electronics in Agriculture, 145*, 311-318.