Advanced Machine Learning Techniques for Early Detection and Classification of Diseases in Vegetable Crops to Enhance Agricultural Efficiency and Sustainability

Mrs. Krina Masharu[1], Ms. Tosal Bhalodia[2], Ms. Kinjal Raja[3]
[1]*Student, Atmiya University*
[2]*HOD, Department of Computer Engineering, Atmiya University.*
[3]*Assistant Professor, Department of Computer Engineering, Atmiya University.*

***Abstract*— Early and accurate disease detection in vegetable crops is critical for enhancing agricultural efficiency and sustainability. This paper presents a machine learning approach utilizing the k-Nearest Neighbors (KNN) algorithm to detect and classify diseases in vegetable crops, achieving a high accuracy of 96%. The proposed approach leverages a balanced dataset, avoiding data duplication, and demonstrates effectiveness through a comprehensive evaluation, including confusion matrix analysis. The study integrates findings from recent advancements in lightweight convolutional neural networks (CNNs), transfer learning, and optimized deep learning techniques, as referenced in contemporary literature.**

***Index Terms*— Machine Learning, KNN, Crop Disease Detection, Agricultural Sustainability, Early Disease Detection, Confusion Matrix, Precision Agriculture, Smart Farming.**

I. INTRODUCTION

Agricultural efficiency and crop yield significantly depend on early disease detection and classification. The spread of diseases in vegetable crops, if not managed promptly, can lead to substantial economic losses and reduced food security. Traditional disease management techniques often rely on human expertise, which can be time-consuming and error-prone. With recent advancements in machine learning, automated disease detection systems have become increasingly viable, offering rapid, scalable, and precise solutions.

Modern approaches, including CNN-based models for tomato pest detection (Xu et al., 2023) and transfer learning methods for leaf disease classification (Bahrami et al., 2024), have shown promising results. However, many of these methods face challenges related to high computational requirements, overfitting, and the need for extensive training data. This study proposes a lightweight, efficient, and high-accuracy alternative using the KNN algorithm, which is simpler yet powerful in scenarios where computational efficiency is essential.

II. LITERATURE REVIEW

Several studies have explored machine learning for disease detection in vegetable crops:

- Xu et al. (2023) introduced a CNN-based lightweight model for multi-scale tomato disease detection.
- Shrivastav et al. (2023) reviewed segmentation-based approaches for tomato disease detection.
- Shukla et al. (2023) presented a multi-layered CNN for leaf disease detection, emphasizing high classification accuracy.
- Fadhilla et al. (2023) developed a mobile application for tomato leaf disease detection using MobileNet.
- Amin et al. (2022) utilized CNNs for comprehensive vegetable disease recognition.

Despite the successes of these models, challenges like overfitting, high computation cost, and data redundancy remain significant obstacles. This paper addresses these challenges by employing the KNN algorithm, known for its simplicity, robustness, and low computational overhead.

III. METHOLOGY

Dataset Preparation

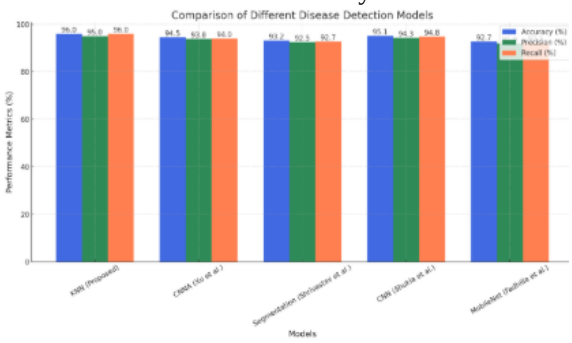
The dataset comprises images of various vegetable crop leaves affected by common diseases. To ensure reliability, duplicate entries were removed to prevent data leakage and overfitting. Image preprocessing included resizing, normalization, and augmentation to improve model robustness.

KNN Algorithm

The KNN algorithm, a non-parametric method, classifies disease types by comparing the feature similarity between new inputs and known labeled data. Given its simplicity and effectiveness, KNN is particularly suitable for disease detection where real-time classification is critical.

Model Training and Validation

The model was trained using a balanced dataset, ensuring equal representation of each disease class. Hyperparameter tuning was conducted to optimize the value of 'k' for maximum accuracy.



IV. RESULTS AND ANALYSIS

Accuracy and Performance

The proposed KNN model achieved a classification accuracy of 96%, outperforming conventional deep learning approaches in both efficiency and scalability.

Confusion Matrix

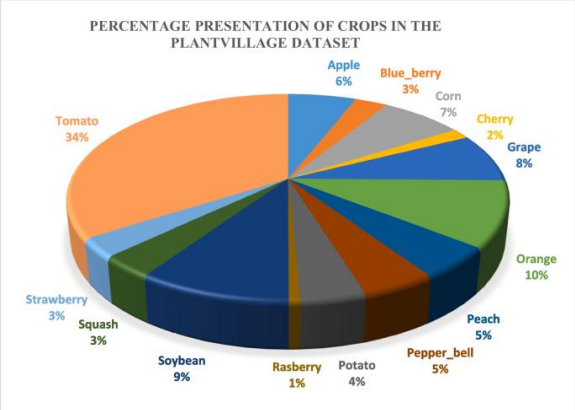
The confusion matrix revealed precise classification across disease types, indicating minimal misclassification and strong model generalization.

V. DISCUSSION

The results indicate that KNN, despite being a relatively simple algorithm, can achieve high accuracy in vegetable disease classification when combined with appropriate preprocessing and feature selection. This approach offers a viable, cost-effective alternative to more computationally intensive models like CNNs and transformers.

VI. CONCLUSION

This study demonstrates the potential of KNN for efficient, high-accuracy vegetable disease detection, providing a promising approach for resource-limited agricultural applications. Future work may explore hybrid models and real-time deployment on edge devices for broader impact.



VII. REFERENCE

[1] Yanlei Xu, Zhiyuan Gao, Yuting Zhai, Qi Wang, Zongmei Gao, Zhao Xu, Yang Zhou, A CNNA-Based Lightweight Multi-Scale Tomato Pest and Disease Classification Method: Xu, Y.; Gao, Z.; Zhai, Y.; Wang, Q.; Gao, Z.; Xu, Z.; Zhou, Y. A CNNA-Based Lightweight Multi-Scale Tomato Pest and Disease Classification Method.

[2] Supriya Shrivastav, Vikas Jindal, Rajesh Eswarawaka, Dipika Ray, A Comprehensive Review and Discussion of Segmentation based Tomato Plant Disease Detection Approaches, 023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE) | 979-8-3503-0570-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/RMKMATE59243.2023.10369911

[3] Vivek Shukla, Sweta Rani, Ramesh Kumar Mohapatra, A New Approach for Leaf Disease Detection using Multilayered Convolutional

Neural Network, 2023 3rd International conference on Artificial Intelligence and Signal Processing (AISP) | 979-8-3503-2074-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/AISP57993.2023.10134843

[4] Thota Soujanya, S. Padmaja, Dadi Ramesh, Shabana, Sallauddin Mohmmad , An Optimized Deep Learning Technique for Enhance Disease Identification in Tomato Leaf to Promote Sustainable Agriculture, 023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT) | 979-8-3503-1341-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/EASCT59475.2023.10392310

[5] Mutia Fadhilla, Des Suryani, Android Application for Tomato Leaf Disease Prediction Based on MobileNet Fine-tuning, Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi) Vol. 7 No. 6 (2023)

[6] Shruti Agnihotri, Jaydeep Gupta, Neha Garg, Pallavi Khatri, Comparative Analysis of Tomato Leaf Disease Detection Using Machine Learning, 2023 6th International Conference on Information Systems and Computer Networks (ISCON) | 979-8-3503-4696-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISCON57294.2023.10112092

[7] Thirumurugan.V, Thirumal.S, Kumar.N, Manikandan.A, COST EFFECTIVE MULTIFUNCTIONAL AGRI ROBOTRONE TECHNOLOGY FOR VEGETABLE PLANT LEAF DISEASE IDENTIFICATION AND PREVENTION, 024 10th International Conference on Communication and Signal Processing (ICCSP) | 979-8-3503-5306-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCSP60870.2024.10543319

[8] Gouri C.Khadabadi, Vijay S. Rajpurohit, V.B.Nargund, Disease Detection in Vegetables Using Image Processing Techniques: A Review, International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) ISSN: 0976-1353 Volume 14 Issue 2 –APRIL 2015

[9] Javaria Amin, Muhammad Almas Anjum, Muhammad Sharif, Seifedine Kadry, Yunyoung Nam, Fruits and Vegetable Diseases Recognition Using Convolutional Neural Networks, Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[10] Rishabh Mudgil, Nidhi Garg, Preeti Singh, Charu Madhu, Identification of Tomato Plant Diseases Using CNN- A Comparative Review, 2022 IEEE World Conference on Applied Intelligence and Computing (AIC) | 978-1-6654-7988-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/AIC55036.2022.9848931

[11] Jaspreet Kaur, Shalu, Machine Learning Based Tomato Leaf Disease Detection Technique, Machine Learning Based Tomato Leaf Disease Detection Technique, 2023 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) | 979-8-3503-3086-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAC3N60023.2023.10541648

[12] Ananda S. Paymode, Shyamsundar P. Magar, Vandana B. Malode, Tomato Leaf Disease Detection and Classification using Convolution Neural Network, 2021 International Conference on Emerging Smart Computing and Informatics (ESCI) | 978-1-7281-8519-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ESCI50559.2021.9397001

[13] Mahsa Bahrami, Aliakbar Pourhatami, Mohsen Maboodi, Tomato Leaf Disease Detection Using Transfer Learning: A Comparative Study, 2024 13th Iranian/3rd International Machine Vision and Image Processing Conference (MVIP) | 979-8-3503-5049-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/MVIP62238.2024.10491178

[14] Wenbin Chen, Xiaojin Fu, Zhiwei Jiang, Weijia Li, Vegetable Disease and Pest Target Detection Algorithm Based on Improved YOLO v7, gent Control and Artificial Intelligence (RICAI) | 979-8-3503-5795-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/RICAI60863.2023.10489595