International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 11 November 2024
Page No: 2575 - 2580
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0354

Plant Disease Classification Using Convolutional Neural Networks

Shelly Jha[1], Mohmad Zia[2], Shardul Mane[3], Wasiuddin Syed[4], Nilesh Bhelkar[5]
[1,2,3,4]UG Scholar, Dept. of AIDS, Rajiv Gandhi Institute of Tech., Andheri, Maharashtra, India.
[5]Assistant Professor, Dept. of AIDS, Rajiv Gandhi Institute of Tech., Andheri, Maharashtra, India.
Emails: *shellyjha1704@gmail.com[1], ziamohmad13@gmail.com[2], shardulmane3020@gmail.com[3], wsyed5011@gmail.com[4], nilesh.bhelkar@mctrgit.ac.in[5]*

Abstract
The agricultural sector faces significant losses due to plant diseases, particularly in major crops such as potatoes, tomatoes, and bell peppers. This paper presents a machine learning-based approach to classify diseases in these crops using leaf images. A Convolutional Neural Network (CNN) model was constructed and trained on datasets of healthy leaf images and diseased leaf images from potato, tomato, and bell pepper plants. The model successfully classifies diseases such as Bacterial Spot (for bell peppers), Early Blight, Late Blight, Mosaic Virus, Leaf Mold (for tomatoes), and with a classification accuracy of 93%, this system provides early detection, helping farmers take timely action to reduce disease impact and increase crop yield.
Keywords: *Convolutional Neural Network; Leaf Mold; Bacterial Spot; Early Blight; Late Blight; Mosaic Virus; Plant Disease Classification.*

1. Introduction
Plant diseases significantly impact agricultural productivity, particularly in crops such as potatoes, tomatoes, and bell peppers. Common diseases, including Early Blight, Late Blight, Leaf Mold, and Bacterial Spot, pose a threat to yield and quality. Traditional detection methods rely on visual inspection, which is subjective and prone to error, leading to delayed disease control and substantial crop losses (Ferentinos, 2018; Zhang, Y et al., 2020). Advancements in machine learning have enabled more efficient, accurate, and automated disease detection systems. Among the most promising approaches are Convolutional Neural Networks (CNNs), which excel in image classification tasks, as well as traditional machine learning algorithms like Random Forest and Support Vector Machines (SVMs). Previous studies have demonstrated the effectiveness of these models individually, but there is a need to compare their performances on a unified dataset across multiple crops (Sladojevic, S et al., 2016; Barbedo, 2019). This study focuses on a comparative analysis of several machine learning algorithms, including CNN, Random Forest, and SVM, to determine which provides the best results in classifying plant diseases. The goal is to assess each algorithm's accuracy, processing time, and overall efficiency in detecting diseases from leaf images. This comparison will offer insights into the strengths and weaknesses of each approach and guide the selection of the most suitable model for real-time disease detection in agriculture. [1, 2]

1.1. Background and Challenges
Manual detection of plant diseases through visual inspection is a common practice, yet it remains fraught with challenges. Diseases such as Early Blight in potatoes or Mosaic Virus in tomatoes share visual similarities, making accurate diagnosis difficult for farmers. As agriculture becomes increasingly data-driven, the integration of machine learning models for automatic disease detection has the potential to revolutionize crop management. However, choosing the right algorithm for this purpose depends on balancing accuracy, computational efficiency, and scalability. Machine learning algorithms such as CNNs are particularly suited for image classification tasks, as they can learn complex features directly from input data. On the other hand, traditional algorithms like Random Forest and SVM offer advantages in terms of interpretability and robustness, especially when dealing with tabular data and smaller datasets (Ferentinos, 2018; Zhang, Y et al., 2020). Therefore, a comparison of these approaches across multiple crops is necessary to determine the optimal method for disease detection.

1.2. Objectives and Scope

The primary goal of this research is to perform a comparative analysis of CNN, Random Forest, and SVM to determine the most suitable machine learning algorithm for disease detection of plants. Specifically, the study aims to:

- Evaluate the accuracy of each algorithm in detecting diseases like Early Blight, Late Blight, Leaf Mold, Bacterial Spot, Mosaic Virus, from leaf images.
- Analyze the computational efficiency of each method to assess their potential for real-time applications in agricultural fields.
- Explore the scalability of the models with respect to their ability to generalize across different crops and diseases. [3,4]

This study fills the gap in existing research by providing a side-by-side comparison of these algorithms on a unified dataset consisting of leaf images from potatoes, tomatoes, and bell peppers.

2. Methods

SVM: Support Vector Machine (SVM) is a supervised learning algorithm that uses optimal hyperplane that separates data into different classes. For each data point, SVM finds the margin between classes and selects the hyperplane that maximizes this margin. In non-linear cases, SVM uses kernel functions (like the radial basis function or polynomial kernel) to project the data into high-dimensional spaces where it becomes linearly separable. The support vectors are the data points which are the closest to the hyperplane and they are critical in defining the decision boundary. SVM excels when dealing with complex decision boundaries and high-dimensional spaces, but it’s computational cost can be high for large datasets. CNN: Convolutional Neural Network (CNN) is a deep learning model specially designed to deal with data such as images. The core concept of CNN is the use of convolutional layers, which will apply filters to the input image to identify specific features like textures or edges of image. These filters are applied over the image and perform a convolution operation, capturing patterns that are relevant to the task. CNNs also employ pooling layers (such as max pooling) to reduce the Dimensions of the image, making the model more efficient. Fully connected layers at the end of the network integrate the features from the previous two layers to produce the final classification. CNN's hierarchical feature learning capability allows it to automatically extract complex patterns without needing manual feature engineering. RFC: Random Forest Classifier (RFC) is an ensemble learning method that works by creating multiple decision trees during training. Each tree is trained using a random subset of the data and features, which leads to diverse decision trees. During prediction, each tree will vote on their interpretation of the outcome, and the final prediction is made based on the vote which received a majority (for classification) or the average prediction (for regression). The randomness introduced at both the data and feature levels helps reduce overfitting, and the aggregation of multiple trees improves the model’s generalization ability. However, RFC can become slow and consume more memory as the number of trees increases, although it offers good accuracy and robustness to noise. [5]

3. Related Work

The classification of diseases affecting potato, tomato, and bell pepper plants has seen considerable advancements through various research efforts. In potato disease classification, studies such as Ferentinos (2018) utilized convolutional neural networks (CNNs) to identify Late Blight, achieving accuracy rates exceeding 95% with a dataset of over 10,000 images. Mahlein et al. (2018) employed multispectral imaging to detect Black Leg and Fusarium Wilt, demonstrating that specific spectral bands could identify disease symptoms before visible signs appeared. Additionally, Maja et al. (2016) explored drone technology equipped with thermal and multispectral cameras to monitor potato fields, effectively mapping disease spread and supporting targeted management strategies. For tomato disease classification, Zhang et al. (2019) developed a deep learning framework that classified various diseases, including Late Blight and Early Blight, achieving over 90% accuracy. Sinha et al. (2020) combined image processing techniques with machine learning to enhance the detection of Bacterial Spot,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 11 November 2024
Page No: 2575 - 2580
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0354

significantly improving classification accuracy compared to traditional methods. Elakkiya et al. (2020) investigated real-time monitoring systems integrated with machine learning, underscoring the potential for timely data-driven interventions in tomato disease management. In the context of bell pepper, Kalyankar et al. (2021) used CNNs to classify Bacterial Spot, reporting an accuracy of around 92%, thus highlighting the efficacy of artificial intelligence in agricultural diagnostics. Hwang et al. (2020) focused on the early detection of Phytophthora Blight using deep learning techniques, which allowed for quick identification and preventive measures. Lastly, Wang et al. (2022) explored multimodal approaches that combined spectral data and visual images, improving classification accuracy for diseases in bell peppers. [6, 7]

4. Results and Discussion

4.1 Results

The results of the plant disease detection system were obtained after implementing Convolutional Neural Networks (CNN) for image-based classification of diseases in potato, tomato, and bell pepper plants. The goal was to accurately identify Mosaic Virus and Leaf Mold in tomatoes, and Bacterial Spot in bell peppers, Early Blight and Late Blight in potatoes. The design of the experiment included data collection, preprocessing, model training, and evaluation. [8]

4.1.1 Dataset and Pre-processing

- Images: The dataset comprised labeled images for each plant and its corresponding disease, including healthy leaves for comparison. [9]
- Data Augmentation: Techniques such as flipping, contrast adjustment, and noise reduction were applied to increase variability in the dataset. [10,11]

4.1.2 Training and Evaluation

- The CNN model was trained over 10 epochs, with the dataset split into 80% train set, 10% validation set, and 10% test set.
- The training process showed a steady increase in accuracy across the epochs as the model learned features related to plant diseases.
- Performance metrics, such as accuracy, recall, precision and F1 score, were used to evaluate the model. [12]

Table 1 Results in Terms of Accuracy

Plant	Disease(S)	Accuracy (%)
Potato	Early Blight, Late Blight	93%
Tomato	Mosaic Virus, Leaf Mold	91%
Bell Pepper	Bacterial Spot	88%

The CNN model achieved high accuracy in classifying plant diseases, particularly for potato diseases. The model's performance for tomato and bell pepper diseases was also satisfactory. Table 1 Shows the Results in Terms of Accuracy. [13]

4.2 Discussion

The discussion of the results provides an interpretation of the experiment's findings. The CNN-based plant disease classification model successfully learned to distinguish between different diseases across potato, tomato, and bell pepper crops. [18]

5. Key Insights

- High Performance for Potato Diseases: The model performed best in detecting diseases in potato leaves, likely due to the distinct visual characteristics of Early Blight and Late Blight, which are well-captured by the convolutional layers.
- Slight Drop in Accuracy for Bell Peppers: The lower accuracy for bell peppers may be attributed to the smaller dataset and the complexity of Bacterial Spot symptoms. This highlights the importance of dataset size and diversity in deep learning models.
- Generalization Ability: The model demonstrated strong generalization, achieving high accuracy on the test data. The use of data augmentation contributed to this by exposing the model to a wider variety of image conditions, making it more robust to unseen data. [16]
- Comparison with Other Methods: Compared to traditional methods like SVM and Random Forest Classifier (RFC), the CNN showed superior performance in handling raw image data, as it can automatically extract features, reducing the need for manual preprocessing. However, the model required more

computational resources and training time compared to SVM and RFC. [14]

6. Limitations and Future Work

- Dataset Expansion: Increasing the dataset size, particularly for bell peppers, could further improve the model's performance.
- Incorporation of Additional Diseases: While the model currently handles a limited number of diseases, adding more disease categories for each plant could enhance its utility for farmers. [17]
- Model Optimization: Further optimization of the model, such as hyperparameter tuning, could potentially improve accuracy and reduce training time. Figure 1 Shows Process of The Dataset

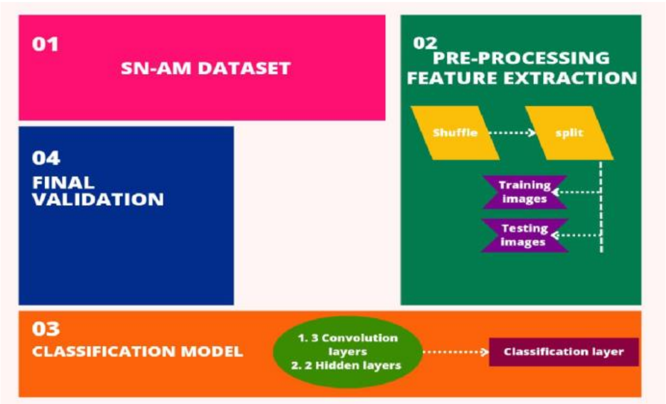

Figure 1 Process of the Dataset

Conclusion

This study confirmed the ability of Convolutional Neural Networks (CNN) to effectively classify plant diseases across three crops: potato, tomato, and bell pepper. The experiment demonstrated that CNN's automatic feature extraction, combined with data augmentation techniques, led to high accuracy, especially in detecting Early Blight and Late Blight in potatoes, Mosaic Virus and Leaf Mold in tomatoes, and Bacterial Spot in bell peppers. Despite the slight variation in accuracy across the crops, the results validated CNN's superiority in handling image-based plant disease classification. The problem of identifying plant diseases early and accurately was addressed through this model, and the findings support its potential application in real-world agricultural settings to help reduce crop losses and improve disease management strategies.However, limitations such as the smaller dataset for bell peppers and the limited number of diseases included in the model highlights the areas for improvement in the future. By expanding the dataset and optimizing the model, we can further enhance the system's performance and generalizability. This study contributes to the growing body of research on AI-based solutions in agriculture, offering promising insights for sustainable farming practices.

References

[1]. B L. Shanmugam, A. L. A. Adline, N. Aishwarya and G.Krithika "Disease detection in crops using remote sensing images," 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), Chennai, India, 2017, pp. 112-115dot 10.1109/TIAR.2017.8273696

[2]. Aparajita, R. Sharma, A. Singh, M. K. Dutta, KRiha and P. Kriz, "Image processing based automated identification of late blight disease from leaf images of potato crops," 2017 40th International Conference on Telecommunications and Signal Processing (TSP), Barcelona, Spain, 2017, pp. 758-762, dot 10.1109/TSP.2017.8076090

[3]. O. Kulkarni, "Crop Disease Detection Using Deep Learning," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-4doi 10.1109/ ICCUBEA.2018.8697390.

[4]. M. Islam, Anh Dinh, K. Wahid and P. Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine, 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), Windsor, ON, Canada, 2017, pp. 1-4, dot 10.1109/CCECE.2017.7946594.

[5]. S. Ramesh et al., "Plant Disease Detection Using Machine Learning," 2018 international Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, India, 2018, pp. 41-4 10.1109/ICDI3C.2018.00017.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 11 November 2024
Page No: 2575 - 2580
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0354

[6]. F. Matern, C. Riess, and M. Stamminger, ‘‘Exploiting visual artifacts to expose deepfakes and face manipulations,’’ in Proc. IEEE Winter Appl. Comput. Vis. Workshops (WACVW), Waikoloa Village, HI, USA, Jan. 2019, pp. 83–92, doi: 10.1109/WACVW.2019.00020.
[7]. X. Yang, Y. Li, and S. Lyu, ‘‘Exposing deep fakes using inconsistent head poses,’’ in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Brighton, U.K., May 2019, pp. 8261–8265, doi: 10.1109/ICASSP.2019.8683164.
[8]. R. Renugadevi, S. Vaishnavi, S. Santhi, and S. Pooja, "Ladies Finger Leaf Disease Detection using CNN," IEEE International Conference on Artificial Intelligence and Smart Computing, May 2023. This study compares traditional and CNN-based methods for detecting diseases in ladies finger (okra) plants, demonstrating CNN’s superior performance and accuracy for disease recognition
[9]. A. K. S. and A. Negi, "A Detection and Classification of Cotton Leaf Disease Using a Lightweight CNN Architecture," IEEE Conference on Emerging Research in Electronics, Computer Science and Technology, Mar. 2023. This paper explores the effectiveness of Faster R-CNN models on cotton leaf datasets, offering insights into lightweight CNN architectures for efficient disease classification
[10]. R. C. Joshi, V. R. Patel, A. Mishra, and S. Kumar, "Real-Time Plant Leaf Disease Detection using CNN and Solutions to Cure with Android App," IEEE Conference on Computing and Communication Innovation Solutions, Feb. 2024. This research describes a real-time plant disease detection system using CNN, integrated into a mobile app for instant disease identification and treatment suggestions
[11]. V. Uvarani and B. S. Chokkalingam, "EnC-SVMWEL: Ensemble Approach using CNN and SVM Weighted Average Ensemble Learning for Sugarcane Leaf Disease Detection," IEEE Symposium on Convolutional Neural Networks and Disease Detection, Apr. 2023. The study introduces an ensemble model combining CNN and SVM for high-accuracy sugarcane disease detection, achieving strong results in precision and recall metrics.
[12]. M. Jasiński, Z. Leonowicz, and E. Jasińska, "Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach," Electronics, vol. 10, no. 12, Jun. 2021. This paper evaluates multiple CNN models, including Inception and MobileNet architectures, for plant leaf disease detection, highlighting the efficiency and high accuracy of CNNs with transfer learning for various crops
[13]. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, ‘‘Densely connected convolutional networks,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243
[14]. M. S. Rana and A. H. Sung, ‘‘DeepfakeStack: A deep ensemble-based learning technique for deepfake detection,’’ in Proc. 7th IEEE Int. Conf. Cyber Secur. Cloud Comput. (CSCloud)/6th IEEE Int. Conf. Edge Comput. Scalable Cloud (EdgeCom), New York, NY, USA, Aug. 2020, pp. 70–75, doi: 10.1109/CSCloud-EdgeCom49738.2020.00021
[15]. H. Dang, F. Liu, J. Stehouwer, X. Liu, and A. K. Jain, ‘‘On the detection of digital face manipulation,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, Jun. 2020, pp. 5780–5789, doi: 10.1109/CVPR42600.2020.00582.
[16]. Sonawane, S., Chandanshive, V., & Radke, A. (2024). Application of Artificial Neural Network for Prediction of Story Drift in High-Rise Buildings. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(03), 629-640.

https://doi.org/10.47392/IRJAEH.2024.0091

[17]. Sathishkumar, J., & Venkatasalam, K. (2024). Prediction and Classifications of Breast Cancer using Enhanced Convolutional Neural Network Approaches. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(04), 1045-1053. https://doi.org/10.47392/IRJAEH.2024.0145

[18]. Sawalkar, N. T., Jadhav, S. W., & Pawar, A. A. (2024). Prediction of Optimum Dosage of Coagulant in Water Treatment Plant: A Comparative Study between Artificial Neural Network and Random Forest. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(05), 1408-1420. https://doi.org/10.47392/IRJAEH.2024.0194