International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Plant Disease Detection Using Machine Learning

Mr. Praveen S R[1], Harshitha S[2], Lakshmi N[3], Shalini G[4]

Assistant Professor[1], Student [2], Student [3], Student [4]

Computer Science and Engineering, R. L. Jalappa Institute of Technology, Doddaballapura, Bengaluru, India.

ABSTRACT—

Agriculture plays a crucial role in the Indian economy. Early detection of plant diseases is very much essential to prevent crop loss and further spread of diseases. Most plants such as apple, tomato, cherry, grapes show visible symptoms of the disease on the leaf. These visible patterns can be identified to correctly predict the disease and take early actions to prevent it. This can be overcome by the use of machine learning and deep learning algorithms. Hence, we are proposing a method that which is detecting the disease of a tomato plant from their leaf images. Here the process is performed with the Support Vector Machine (SVM), Random Forest algorithms of machine learning along with deep learning algorithms Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and ResNet, which is a one of the transfer learning method of CNN. Once after training the dataset with the algorithms, the accuracy of algorithms is compared and the images are classified. And the precautions are also provided for the classified plant. Early identification of plant diseases is crucial for sustainable agriculture. This study presents an innovative machine learning framework for automated plant disease detection using convolutional neural networks (CNNs). The system integrates image segmentation techniques with deep learning algorithms to extract distinctive disease features while minimizing false positives. Results demonstrate superior performance compared to traditional detection methods, with significantly reduced diagnostic time.

I. INTRODUCTION

The global agricultural sector faces a growing challenge in detecting and managing plant diseases.As the world population grows, ensuring food security through healthy crop yields becomes critical. Traditional plant disease detection methods are often slow, labor-intensive, and require expert knowledge. Thus, there is a pressing need for automated, accurate systems that can detect plant diseases efficiently and provide timely interventions to mitigate crop damage. Plant diseases are a significant challenge in agriculture that can adversely impact crop production and quality, leading to economic losses for farmers and threatening food security for populations. Plant diseases are caused by a range of pathogens such as viruses, bacteria, fungi, and nematodes, which can infect plant tissues, affect plant growth and development, and cause symptoms such as leaf spots, wilting, stunting, and discoloration. Identifying and diagnosing plant diseases accurately and quickly is critical for farmers to take appropriate measures to control their spread and limit crop losses. Agriculture is an important human effort that contributes significantly to a country's economic prosperity. Yet, additional challenges have emerged, making it increasingly difficult to sustain crop production and provide food security.

These difficulties include the growing global population, the effects of climate change, a scarcity of cultivable land, and the frequency of plant diseases. Because of the world's fast population growth, the agriculture business is experiencing increasing pressures. This expansion has boosted demand for food production, putting enormous pressure on the industry to produce more crops. Sadly, the continuous effects of climate change have made this effort more difficult, resulting in unpredictable weather patterns and erratic rainfall, making it harder to cultivate crops successfully. Another problem affecting agriculture is a scarcity of arable land. Farmers are driven to utilize more land when cultivable land becomes scarce, resulting in unsustainable land use practices. This, in turn, has a detrimental influence on the environment and the agriculture industry's long- term survival. Moreover, plant diseases, pests, and other biotic factors have led to severe crop losses, lowering food output even further. Solving these issues would necessitate a multifaceted approach that takes into consideration the need to boost food production while also maintaining sustainability and environmental preservation. It is critical to provide creative solutions that increase agricultural yields, reduce land use, and efficiently manage pests and diseases. Only by taking such actions will the agriculture industry be able to continue contributing to economic growth while satisfying the needs of a growing global population.

II. PROBLEM STATEMENT

Plant diseases pose a substantial threat to global agriculture, resulting in significant crop damage and economic losses. Traditional methods of disease detection are often time-consuming and require expertise, making early intervention difficult. This research addresses the need for an automated, reliable, and efficient plant disease detection system using machine learning and deep learning algorithms. The primary challenge is to develop a model capable of accurately identifying diseases from images of plant leaves, thereby enabling timely interventions to mitigate crop.

The problem of plant disease detection using machine learning encompasses multiple technical, practical, and methodological challenges that must be addressed to develop effective and deployable systems. The fundamental problem lies in creating automated systems that can accurately identify and classify plant diseases from visual symptoms while overcoming the inherent complexities and variabilities present in agricultural environments.

The primary technical challenge involves the complexity of visual disease symptoms, which can vary significantly based on disease stage, environmental conditions, plant variety, and co-occurring stresses. Many diseases present similar visual symptoms, particularly in early stages, making accurate classification difficult. Symptoms may overlap between different diseases, and plants often suffer from multiple diseases simultaneously, complicating the diagnostic process. Additionally, disease symptoms can be confused with nutrient deficiencies, environmental stress, or pest damage, requiring sophisticated classification algorithms to distinguish between these conditions. Data-related problems present significant challenges in developing robust machine learning models. High-quality, labeled datasets for plant diseases are often limited, particularly for less common diseases or crop varieties. The collection of representative datasets requires expertise in plant pathology and significant time investment. Image datasets must capture the full spectrum of disease progression, various environmental conditions, and different plant growth stages to ensure model generalizability. The imbalanced nature of disease datasets, where some diseases are well-represented while others have few examples, poses additional challenges for training effective classification models.

Environmental variability represents another major problem, as field conditions vary dramatically in terms of lighting, weather, background interference, and image quality. Machine learning models must be robust to these variations while maintaining diagnostic accuracy. The problem extends to seasonal variations, geographical differences, and the dynamic nature of agricultural environments where conditions change rapidly. Deployment challenges involve translating laboratory- developed models into practical field applications. This includes addressing computational constraints of mobile devices, ensuring reliable performance in areas with limited internet connectivity, and creating interfaces that can be effectively used by farmers with diverse educational backgrounds and technical expertise. The integration of machine learning systems with existing agricultural practices and decision-making processes presents additional implementation challenges.The problem also encompasses the need for continuous model improvement and adaptation as new diseases emerge, existing diseases evolve, and crop varieties change. This requires developing systems capable of incremental learning and adaptation while maintaining reliability and accuracy in disease detection and classification**.**

III. LITERATURE REVIEW

In 2015, S. Khirade et Al. tackled the problem of plant disease detection using digital image processing techniques and back propagation neural network (BPNN) [1]. Authors have elaborated different techniques for the detection of plant disease using the images of leaves. They have implemented Otsu's thresholding followed by boundary detection and spot detection algorithm to segment the infected part in leaf. After that they have extracted the features such as color, texture, morphology, edges etc. for classification of plant disease. BPNN is used for classification i.e. to detect the plant disease. Shiroop Madiwalar and Medha Wyawahare analyzed different image processing approaches for plant disease detection in their research [2]. Authors analyzed the color and texture features for the detection of plant disease. They have experimented their algorithms on the dataset of 110 RGB images. The features extracted for classification were mean and standard deviation of RGB and YCbCr channels, grey level cooccurrence matrix (GLCM) features, the mean and standard deviation of the image convolved with Gabor filter. Support vector machine classifier was used for classification. Authors concluded that GCLM features are effective to detect normal leaves. Whereas color features and Gabor filter features are considered as best for detecting anthracnose affected leaves and leaf spot respectively. They have achieved highest accuracy of 83.34% using all the extracted features. Peyman Moghadam et Al. demonstrated the application of hyperspectral imaging in plant disease detection task [3]. visible and near-infrared (VNIR) and short-wave infrared (SWIR) spectrums were used in this research. Authors have used k-means clustering algorithm in spectral domain for the segmentation of leaf. They have proposed a novel grid removal algorithm to remove the grid from hyperspectral images. Authors have achieved the accuracy of 83% with vegetation indices in VNIR spectral range and 93% accuracy with full spectrum. Though the proposed method achieved higher accuracy, it requires the hyperspectral camera with 324 spectral bands so the solution becomes too costly.

[4] S. S. Sannakki and V. S. Rajpurohit, proposed a "Classification of Pomegranate Diseases Based on Back Propagation Neural Network" which mainly works on the method of Segment the defected area and color and texture are used as the features. Here they used neural network classifier for the classification. The main advantage is it Converts to L*a*b to extract chromaticity layers of the image and Categorisation is found to be 97.30% accurate. The main disadvantage is that it is used only for the limited crops. [5] P. R. Rothe and R. V. Kshirsagar introduced a" Cotton Leaf Disease Identification using Pattern Recognition Techniques" which Uses snake segmentation, here Hu's moments are used as distinctive attribute. Active contour model used to limit the vitality inside the infection spot, BPNN classifier tackles the numerous class problems. The average classification is found to be 85.52%. [6] Aakanksha Rastogi, Ritika Arora and Shanu Sharma," Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic". K-means clustering used to segment the defected area; GLCM is used for the extraction of texture features, Fuzzy logic is used for disease grading. They used artificial neural network (ANN) as a classifier which mainly helps to check the severity of the diseased leaf.

[7] Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa, proposed" Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease "Color histograms are extracted and transformed from RGB to HSV, RGB to L*a*b.Peak components are used to create max tree, five shape attributes are used and area under the curve analysis is used for classification. They used nearest neighbors, Decision tree, random forest, extremely randomized tree, Naïve bayes and SV classifier. In seven classifiers extremely, randomized trees yield a very high score, provide real time information provide flexibility to the application.

IV. EXISTING SYSTEMS

Traditional Approaches: Current plant disease detection systems primarily rely on visual inspection by agricultural experts, farmers, or extension workers. Many existing automated systems use basic image processing techniques with handcrafted features like color histograms, texture analysis, and shape descriptors. Some systems employ classical machine learning algorithms such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), or decision trees with manually engineered features.

Limitations of Current Systems: These traditional methods often suffer from limited accuracy, especially under varying environmental conditions. They typically require controlled lighting and image acquisition setups, making them impractical for field deployment. The handcrafted features may not capture complex disease patterns effectively, and the systems often work well only for specific crops or disease types they were trained on.

Early ML-Based Systems: Some existing systems have started incorporating machine learning but often use shallow learning approaches. These systems require extensive preprocessing and feature engineering, have limited generalization capabilities, and struggle with complex disease symptoms that may appear similar across different conditions.

V. PROPOSED SYSTEMS

The proposed plant disease detection system employs a comprehensive architecture that combines advanced machine learning techniques with user-friendly interfaces to provide accurate and timely diagnosis of plant diseases. The system is designed to be flexible, scalable, and deployable across various environments, from research laboratories to actual farm fields.

Advantages of Proposed System

1. **Early Detection:** Identifies diseases at an early stage, allowing timely treatment and reducing crop loss.
2. **High Accuracy:** Machine learning and deep learning models (e.g., CNNs) provide precise and consistent results.
3. **Time-Saving:** Automates the process, reducing the need for manual inspection by experts.
4. **Cost-Effective:** Minimizes the cost of disease diagnosis over time, especially in large-scale farming.
5. **User-Friendly:** Can be deployed via mobile apps, making it accessible to farmers with smartphones.
6. **Scalable:** Can be applied across various crops and extended to multiple disease types.
7. **Reduces Chemical Use:** Helps apply pesticides only when necessary, promoting eco- friendly farming.

VI. ARCHITECTURAL MODEL

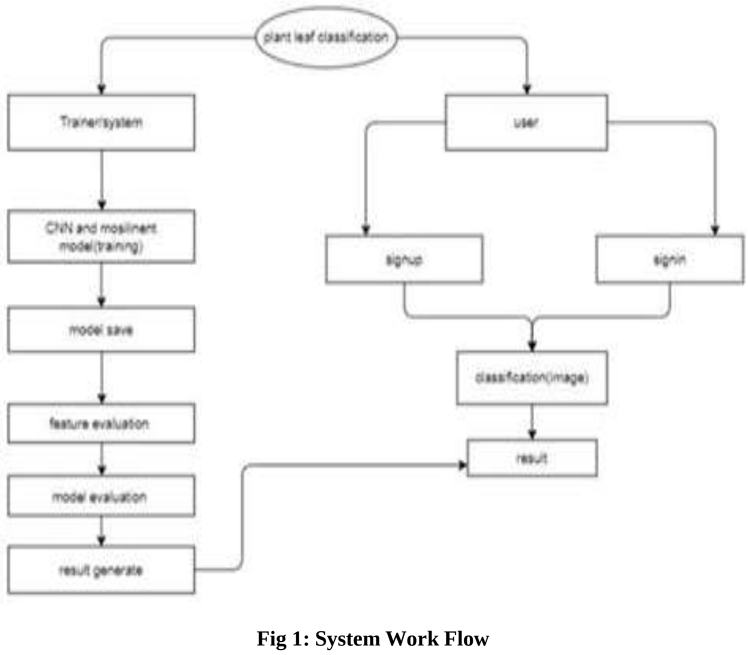

Fig 1: System Work Flow

This diagram illustrates a comprehensive machine learning system for plant disease detection with two main components: a Training System and a User Interface.

Training System

1. **System:** The initial setup phase where the training environment is established
2. **CNN and MobileNet Model Training:** Uses Convolutional Neural Networks (CNN) combined with mobileNet architecture - MobileNet is particularly suitable for plant disease detection as it's lightweight and efficient for mobile deployment while maintaining good accuracy
3. **Model Save:** The trained model is saved for later use
4. **Feature Evaluation:** The system evaluates which features the model has learned to identify (likely disease symptoms, leaf patterns, discoloration, etc.)
5. **Model Evaluation:** Performance metrics are calculated (accuracy, precision, recall, etc.)
6. **Result Generate:** Training results and model performance statistics are generated

User Interface System

1. **User:** End users (farmers, agricultural experts, etc.) access the system
2. **Sign Up/Sign In:** Authentication system to manage user accounts
3. **Classification(image):** Users upload images of plant leaves suspected of disease
4. **Result:** The system provides disease classification results, likely including:Disease type identification,Confidence scores, Plant Type,Health Status

VII. RESULTS

Fig 2: Login Page

Fig 3: To upload a leaf image

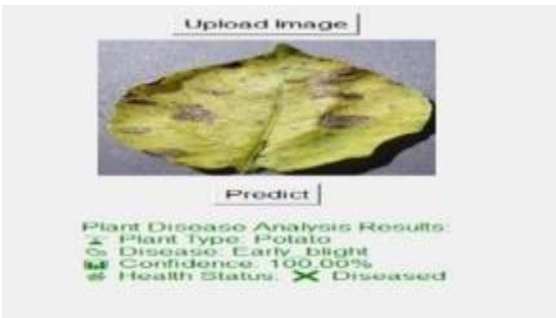

Fig 4: Predicting the results of diseased leaf

Fig 5: Predicting the results of healthy leaf

CONCLUSION

The plant disease detection system developed in this project represents a significant advancements in the application of machine learning to agricultural challenges. By leveraging state-off-the-art computer vision and deep learning techniques, the system provides farmers and agricultural professional with a powerful tool for early disease detection and management. The plant disease detection system developed in this project demonstrates the transformative potential of artificial intelligence in addressing critical agricultural challenges. By making advanced disease detection technology accessible to farmers worldwide, the system contributes to food security, economic sustainability, and environmental protection. The modular and expandable nature of the system ensures it can evolve with advances in machine learning technology and adapt to the changing needs of the agricultural sector. As climate change and globalization continue to alter disease patterns and create new challenges for farmers, tools like this will become increasingly valuable for sustainable and productive agriculture. The success of this project illustrates the power of interdisciplinary collaboration, combining expertise in computer vision, machine learning, plant pathology, and user experience design to create a solution that is both technically sophisticated and practically useful. Moving forward, this approach to agricultural technology development emphasizing both technical excellence and user-centered design will be essential for addressing the complex challenges facing global food production.

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] 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.

[3] Zhang,S.,Wu,X.,&You,Z**.**(2017). Leaf image based cucumber disease recognition using sparse representation classification. Computers and Electronics in Agriculture, **134**, 135–141.

[4] Ferentinos, K.P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311-318.

[5] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018) Deep learning in agriculture, a survey that provides an overview of ML/DL applications in agriculture.

[6] 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.

[7] Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D.P. (2017). Deep learning for image-based cassava disease detection. Frontiers in Plant Science, 8, 1852.

[8] Fuentes, A., Yoon, S., Kim, S.C., & Park, D.S. (2018). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors, 18(9), 2906.

[9] Ferentinos, K. P**.** (2018) uses CNNs for real-time plant disease classification.