Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 46 No. 3 (2025)

Methods for Detecting and Classifying Rice Plant Diseases using Machine Learning Technique

Metharani N [1], Manoranjitham T [2]

[1, 2] *Department of Computing Technologies, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu District, Tamil Nadu, India*

Abstract:- Rice is the main source of energy for more than half of the world's population. In order to reduce agricultural loss in the paddy field, this work concentrated on creating a prediction model. First, illnesses of rice plants and their pictures were taken. Next, a massive dataset was encountered using a big data framework. In order to produce the reduced data with significant features that are utilized as the input to the classification model, the feature extraction procedure is applied to the data first, followed by feature selection. A rough set theory-based feature selection method is employed for the feature selection task after features based on color, shape, position, and texture are retrieved from the photos of sick rice plants for the rice disease datasets. In order to create an effective illness prediction model, ensemble classification techniques have been applied to the classification problem within a map-reduce framework. The effectiveness of the suggested model is demonstrated by the outcomes on the gathered disease data.

Keywords*:* Plant disease, machine learning, deep learning, image segmentation.

1. Introduction

A lot of data with a wide range of authenticity can be found in agricultural data. Because new data is often generated from every application and added to the current data, data velocity is likewise very rapid. Managing this type of intricate data is an extremely difficult task in the data mining sector. There are a variety of structured, semi-structured, and unstructured data types; the data are not always structured. Prior to analyzing these data, appropriate structure must be provided using various IT tools and methodologies.

Therefore, a proper preprocessing step is required before using any method. The various kinds of rice disorders and their symptoms are covered in Section 1.1 that follows.

The agriculturist in provincial regions may think that it's hard to differentiate the malady which may be available in their harvests. It's not moderate for them to go to an agribusiness office and discover what the infection may be. Our principal objective is to distinguish the illness introduced in a plant by watching its morphology by picture handling and machine Learning. Pests and Diseases result in the destruction of crops or part of the plant resulting in decreased food production leading to food insecurity.

Also, knowledge about pest management or control and diseases are less in various less developed countries. Toxic pathogens, poor disease control, drastic climate changes are one of the key factors which arise in dwindled food production. Various modern technologies have emerged to minimize postharvest processing, to fortify agricultural sustainability and to maximize productivity.

2. Literature Review

[1] 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 affected area and color and texture are used as the features. Here they used a 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.

[2] 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 a distinctive attribute. Active contour model used to limit the vitality inside the infection spot,

The BPNN classifier tackles numerous class problems. The average classification is found to be 85.52%.

[3] Aakanksha Rastogi, Ritika Arora and Shanu Sharma," Leaf Disease Detection and Grading using Computer Vision Technology & Fuzzy Logic". K-means clustering is used to segment the affected area; GLCM is used for the extraction of texture features, Fuzzy logic is used for disease grading. They used an artificial neural network (ANN) as a classifier which mainly helps to check the severity of the diseased leaf.

[4] 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, randomized trees yield a very high score, provide real time information and provide flexibility to the application.

[5] uan Tian, Chunjiang Zhao, Shenglian Lu and Xinyu Guo," SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases," Color features are represented in RGB to HIS, by using GLCM, seven invariant moment are taken as shape parameter. They used an SVM classifier which has MCS, used for detecting disease in wheat plants offline.

3. Proposed Methodology

A. NB Technique

The Naive Bayes (NB) classifier is a classification algorithm based on the Bayes theorem and the assumption that all predictors are independent of one another. Since this algorithm is based on probabilities, it is necessary to explore the sample distribution and feature type. The NB Technique It is a probabilistic classifier variation built on the NB classifier idea.

B. KNN Technique

The k-nearest neighbors (KNN) algorithm is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. It is one of the popular and simplest classification and regression classifiers used in machine learning today. The KNN Technique It is a nonparametric, supervised ML technique commonly applied to pattern recognition.

C. Decision Tree Technique

The DT Technique In supervised learning, it is a supervised classification and regression algorithm

D. SVM Technique

The SVM Technique The separating hyperplane defines this supervised ML classifier.

E. Random Forest Technique

The RF Technique It is a collection of learning techniques for randomized DT classifiers.

4. Machine Learning Algorithms

Various ML algorithms are employed for classifying diseases, including:

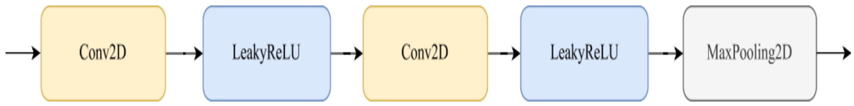

Fig. 1. Machine Learning model concept

ML Technique

1) Support Vector Machines (SVMs)

SVMs are used to create decision boundaries that separate different disease classes, according to a study on Nature.

2) Convolutional Neural Networks (CNNs)

*C*NNs are powerful for learning complex patterns in images, and have been successfully used for rice disease detection and classification.

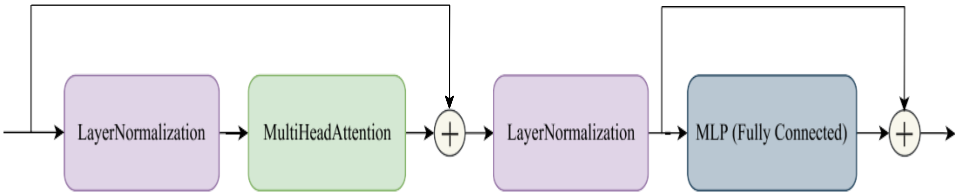

Fig.2. Machine learning workflow

3) Probabilistic Neural Networks (PNNs)

Fuzzy logic, and other algorithms have also been explored for rice disease classification.

4) Deep Learning Approaches

Deep learning models like YOLOv8, VGG16, and ResNet, are used for accurate disease detection and classification.

i. Hybrid Approaches:

Combining different techniques, such as CNNs with other ML algorithms, can lead to improved accuracy and efficiency.

ii. Mobile Applications:

Mobile apps have been developed to facilitate field diagnosis of rice diseases, utilizing techniques like fuzzy entropy and probabilistic neural networks.

Examples of Rice Diseases and Their Symptoms:

iii. Rice Blast: Oval or elliptical spots with reddish-brown margins on the leaves.

iv. Brown Spot: Round or oval dark brown spots on the leaves.

v. Sheath Blight: Irregular spots with greenish-grey or brownish margins.

vi. Bacterial Leaf Blight: Yellowish lesions with uneven edges, leading to yellowing and dying of leaves.

5. Plant Disease Detection

The elongated, yellow sign of bacterial leaf blight is seen on the leaf. Though they differ from one another, the Brown Spot and Leaf Smut are rather similar. In contrast to the Leaf Smut, the Brown Spot produces circles on the leaves that are darker in color and have a larger area. Little spots are dispersed across the leaf in cases of Leaf Smut infection.

Fig. 3. Brown Spot Rice Leaf

Fig. 4. Healthy Rice Leaf

A. Background

Historical detecting method, however beneficial on a plant, presents certain limitations where implemented in extensive agriculture. Initially, physical information to detecting the region of the model laborious in the human resources when agricultural operations are extensive, as there is little margin for mistake in efforts to halt the spread of a disease

B. Image Acquisition

High-resolution images of rice leaves are captured using smartphones, drones, or cameras under natural/artificial lighting. Often labeled with disease types: e.g., Brown Spot, Bacterial Leaf Blight, Leaf Smut, etc.

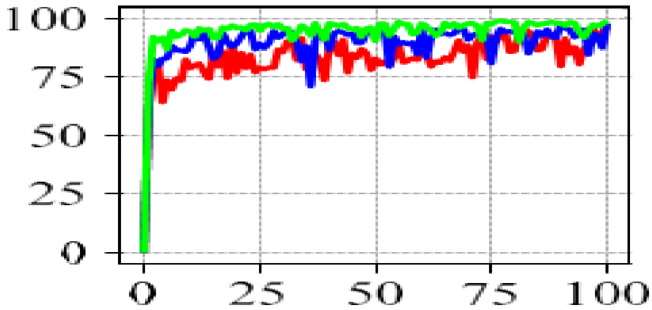

Fig. 5. Plant Disease Detection sample

C. Dataset

This dataset is curated for the purpose of detecting and classifying four distinct types of rice leaf diseases: Hispa, Brown Spot, Leaf Blast, Leaf Scaled, Narraow brown spot and Neck Blast. Each category represents a specific pathology affecting rice plants.

D. Plant Village Dataset

Researchers and enthusiasts in the field of agricultural technology and plant pathology can leverage this dataset to develop robust solutions for early diagnosis and effective management of rice crop health.

E. Rice Leaf Disease Dataset

There are three classes in this data- set: 1-Bacterial Leaf Disease, -Leaf Smut, 2-Brown Spot, and Bacterial Leaf Disease. There are 40 samples in each class. Figure 2 displays three randomly chosen examples from the dataset. As seen in Figures 2a and b, respectively, the background of these photos is either eliminated or replaced with white, or it is quite basic, as shown in Figure 2.

6. Experimental Analysis

To identify plant illnesses, the suggested approach makes use of four different cutting-edge models: CNN, MobileNetV2, EfficientNetB0, and ResNet-50. The findings displayed.

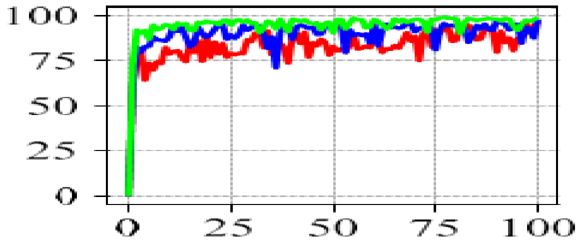

Fig. 6. Disease Sample

Rice plant disease detection using machine learning on Kaggle involves leveraging computer vision techniques, particularly convolutional neural networks (CNNs), to identify different diseases in rice leaves.

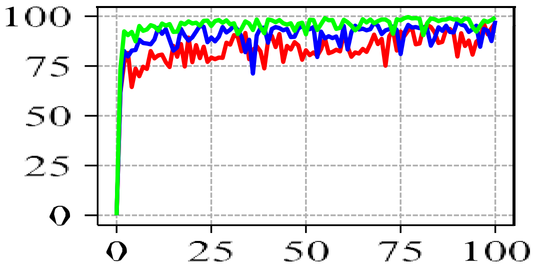

Fig.7. Machine Leaning proposed model

A number of variables, including the crop's economic significance, the disease's severity and prevalence, the profitability of the crop, the the possible yield losses as well as the expense of controlmethods. But in Table 1, the "red" color indicates the illnesses that are most economically significant for each plant, whereas the "green" color indicates healthy examples. This dataset is also unbalanced. It indicates that there are not an equal number of photos in each class.

7. Challenges

The following is a list of several problems and elements that could influence the classification and identification of plant diseases:

1. Dataset of Insufficient Size and Variety
2. Image Segmentation
3. Identification of Diseases with Visually Similar Symptoms
4. Simultaneous Occurrence of Multiple Disease
5. Identification of Diseases from Real-Time Images
6. Design a Light Deep Learning Model

8. Results and Discussion

To identify plant illnesses, the suggested approach makes use of four different cutting-edge models: CNN, MobileNetV2, EfficientNetB0, and ResNet-50. The findings displayed.

Table. 1. Test result for model accuracy

Models	Res	Result of experiments				
		Average F1	Average recall	Average precision	GFlops	Convergence score
Model 1	50	0.97	0.96	0.97	0.006	0.88
	100	0.98	0.98	0.98	0.03	0.93
	200	0.98	0.98	0.99	0.12	0.93

Models	Res	Result of experiments				
		Average F1	Average recall	Average precision	GFlops	Convergence score
Model 2	50	0.98	0.98	0.98	0.007	0.91
	100	0.97	0.97	0.97	0.032	0.91
	200	0.99	0.99	0.99	0.14	0.93
Model 3	50	0.98	0.98	0.98	0.003	0.96
	100	0.99	0.99	0.99	0.012	0.92
	200	**1**	**1**	**1**	**0.053**	**0.97**

Table 1 demonstrates that EfficientNetB0 performs better than four other models in terms of recall, accuracy, and precision. MobileNetV2 is doing well, with the greatest classification performance of 96.89%, after EfficientNetB0. ResNet-50's accuracy rate of 79.83% puts it in last place among its rivals. With the highest accuracy across the models, these findings show how effective the EfficientNetB0 model.

In terms of the quick and precise diagnosis of plant diseases, these technologies can offer models that are on par with, if not better than, human beings. Economic stability, poverty alleviation, food shortages, and unemployment reduction all depend on agriculture. However, diseases and pests have a negative impact on crop yield and seriously damage the economy.

As a result, monitoring makes early detection easier, which is crucial for sustainable resource use and ideal crop management. One section of the peppers was infected with the disease, and the other section was used as a control. The GLCM and LBP techniques were used to analyse the spectral and textural features that were extracted using multispectral imaging. Classification models were created and a variety of feature selection techniques were used.

9. Conclusion

This article compared a convolutional-based architecture of comparable complexity with a simpler version of the new Vision Transformer. Additionally, the hybrid models that combined the CNN and ViT were examined. A small dataset (Rice Leaf Disease Dataset), a medium dataset (Wheat Rust Classification Dataset), and a big dataset (Plant Village) were all used to solve classification problems using these networks. The limited number of samples and the similarities between the two classes were the primary issues with the RLDD dataset.

Rice plant disease detection on Kaggle offers a valuable platform for researchers and enthusiasts to develop and evaluate machine learning models for automated disease diagnosis, ultimately benefiting farmers and improving crop yields. The physical assessment of crops is significantly lacking. Even highly qualified plant specialists may fail to get accurate diagnoses because of the complexity of illnesses, significant deficiency symptoms, or other stress conditions that may present similarly, often exhibiting modest indicators.

Acknowledgement:

The submitting author is gratefully acknowledge the Full time Ph.D. outcome of research work Under Visvesvaraya Ph.D. scheme Phase ll for ESDM/IT/|TE is supported by MietY's Ministry of Electronics and Information Technology (IT/ITES), Government of India, implemented by Digital India Corporation.

References

[1] PeerJ Comput Sci 7:e432. Rahmat B, Purbasari IY, and Putra CS (2021) Convolutional neural networks are used to detect illnesses in rice plants. 1125(1):012021, IOP Conf Ser Mater Sci Eng. 10.1088/1757-899X/1125/1/012021 https://doi.org

[2] Azim, M. A., Islam, M. K., Rahman, Md. M., & Jahan, F. (2021). An effective feature extraction method for rice leaf disease classification. TELKOMNIKA, 19(2), 463. https://doi.org/10.12928/telkomnika.v19i2.16488

[3] Biau, G., & Scornet, E. (2016). A random forest guided tour. TEST, 25(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7
[4] Daniya, T., & Vigneshwari, D. S. (2019). A review on machine learning techniques for rice plant disease detection in agricultural research. International Journal of Advanced Science and Technology, 28(13), 49-62.
[5] FAO. (2018). Asia and the Pacific regional overview of food security and nutrition 2018: Accelerating progress towards the SDGs. Food & Agriculture Organization of the United Nations.
[6] FAO. (2019). Global Report on Food Crises. Food & Agriculture Organization of the United Nations.
[7] Gupta, S., & Sarvaiya, J. N. (2020). Emerging Technology Trends in Electronics, Communication and Networking: Third International Conference, ET2ECN 2020, Surat, India, February 7–8, 2020, Revised Selected Papers (Vol. 1214). Springer Singapore. https://doi.org/10.1007/978-981-15-7219-7
[8] Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. Computers and Electronics in Agriculture, 143, 23-37.
[9] Kemenkes. (2018). Hasil Utama Riskesdas 2018. Badan Penelitian dan Pengembangan Kesehatan-Kementerian Kesehatan.
[10] Li, N., Lv, X., Xu, S., Li, B., & Gu, Y. (2020). An improved water surface images segmentation algorithm based on the otsu method. Journal of Circuits, Systems and Computers, 29(15), 2050251. https://doi.org/10.1142/S0218126620502515
[11] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.
[12] Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In Machine Learning (pp. 101–121). Elsevier. https://doi.org/10.1016/B978-0-12-815739-8.00006-7
[13] Shah, J. P., Prajapati, H. B., & Dabhi, V. K. (2016). A survey on detection and classification of rice plant diseases. 2016 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC), 1–8. https://doi.org/10.1109/ICCTAC.2016.7567333
[14] Swain, S., Nayak, S. K., & Barik, S. S. (2020). A review on plant leaf diseases detection and classification based on machine learning models. Mukt Shabd Journal, 9(7), 5195-5205. https://doi.org/09.0014.MSJ.2020.V9I6.0086781.105023
[15] Tharwat, A. (2016). Linear vs. quadratic discriminant analysis classifier: A tutorial. International Journal of Applied Pattern Recognition, 3(2), 145. https://doi.org/10.1504/IJAPR.2016.079050
[16] Yang, F.-J. (2018). An Implementation of Naive Bayes Classifier. 2018 International Conference on Computational Science and Computational Intelligence (CSCI), 301–306. https://doi.org/10.1109/CSCI46756.2018.00065
[17] Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2020). An overview of blockchain technology: Architecture, consensus, and future trends. In IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) (pp. 1861-1865). IEEE.
[18] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," IEEE Trans. on Neural Networks, vol. 4, pp. 570-578, July 1993.
[19] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan. 1959.
[20] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, "Rotation, scale, and translation resilient public watermarking for images," IEEE Trans. Image Process., vol. 10, no. 5, pp. 767-782, May 2001.