OPEN ACCESS JOURNALS
Indiana Journal of Multidisciplinary Research
Abbreviated Tittle- Ind J Multi Res
ISSN *(Online)*- 2583-3820
Journal Homepage Link- https://indianapublications.com/Journals/IJMR
DOI: https://doi.org/10.5281/zenodo.21600476
INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Research Article
Volume-06|Issue-04|2026

Maize Identification and Categorization of Leaf Diseases Using Machine Learning

Gagana S L[1], Shalini Hanok[2], Ganavi B N[3], Akash L[4], Mrudula M Gowda[5]

[1,2,3,4,5]Department of Electronics and Communication, ATME College of Engineering, Mysore, India.

Article History
Received: 05.05.2026
Accepted: 22.06.2026
Published: 25.07.2026

Citation
Gagana, S. L., Hanok, S., Ganavi, B. N., Akash, L. & Gowda, M. M. (2026). Maize Identification and Categorization of Leaf Diseases Using Machine Learning. *Indiana Journal of Multidisciplinary Research, 6*(4), 164-168.

Abstract: A machine learning-based web application for maize leaf disease classification was developed to facilitate the early detection and classification of diseases affecting maize plants. The system employs the Random Forest algorithm to classify maize leaf images into four categories: Healthy, Gray Leaf Spot, Common Rust, and Blight. The training dataset consists of 1,006 images of Common Rust, 574 images of Gray Leaf Spot, 1,146 images of Blight, and 1,162 images of Healthy leaves, ensuring effective model training and classification. The trained Random Forest model was saved and integrated into a Python Flask-based web application, enabling users to upload images of maize leaves for real-time disease prediction. Upon image upload, the system accurately identifies the disease or confirms the leaf's healthy condition, providing quick and reliable results. This web-based tool offers farmers, agricultural professionals, and researchers an accessible solution for the timely detection and management of maize leaf diseases, thereby supporting improved crop management and disease control. The proposed model achieved an accuracy of 83% during testing, demonstrating its effectiveness for maize leaf disease classification.

Keywords: Maize Leaf Disease, Machine Learning, Random Forest, Gray Leaf Spot, Common Rust, Blight, Disease Detection, Python Flask, Accuracy.

INTRODUCTION

Maize Leaf Disease Classification System was created to help with the early identification inventive machine learning based tool called the ion and control of diseases that impact maize plants. A staple crop in many regions of the world, maize is susceptible to a number of diseases that can drastically lower output and quality. Accurate and timely detection of these illnesses is essential to minimizing their effects and guaranteeing the greatest possible crop yield. This system divides photos of the leaves of maize into four groups: The Gray Leaf Spot, Common Rust, Blight, and Healthy. It doesthis by utilizing the Random Forest algorithm, a powerful machine learning technique. Thousands of photos make up the model's training dataset, which offers a balanced and varied basis for accurate disease diagnosis [1]. The Random Forest algorithm can efficiently differentiate between various kinds of maize leaves by examining their visual characteristics. between sick and healthy leaves, guaranteeing precise forecasts. Users can simply upload photos of maize leaves for examination thanks to the system's integration into a Python Flask web application. The program analyzes the data when a photograph is uploaded [2] and makes a forecast regarding the condition of the leaf or whether a particular illness.

For farmers, agricultural experts, and researchers who require fast and accurate information to manage and treat maize leaf diseases, the system is a useful tool because it provides data instantly. Following the upload of an image, the model analyzes the information and makes a prediction about the health of the leaf or the existence of a certain disease. Because it provides data instantly, the system is a useful resource for farmers, agricultural experts, and researchers who require fast and accurate information in order to manage and control illnesses of the maize leaves [3].Thousands of photos make up the dataset that was used to train the algorithm, offering a varied and balanced basis for accurate disease diagnosis. The The Random Forest algorithm is capable of reliably identify healthy and ill leaves by examining the visual characteristics of maize leaves, guaranteeing precise forecasts. To lessen their effects and guarantee ideal crop yield, these diseases must be identified early and accurately. A web tool based on machine learning, the System of Classification for Maize Leaf Disease was created to help with the early identification and categorization of illnesses in maize plants. The system divides pictures of maize leaves into four categories using the Random Forest method. Using machine learning algorithms [4], this work aims to create a model capable of properly identifying and categorizing leaf diseases at an early stage and taking the appropriateprecautions against them. To identify and categorize diseases from plant leaves, this study employs supervised machine learning techniques classification approaches, including NB, KNN, DT, SVM, and RF. Additionally, a comparison of the various classification strategies is conducted. Additionally, it offers a variety

***Corresponding Author:** Shalini Hanok

of methods that in contrast to other methods, will yield the most accurate results. Numerous applications, including biological signal processing and healthcare, have successfully implemented these classification techniques [4].

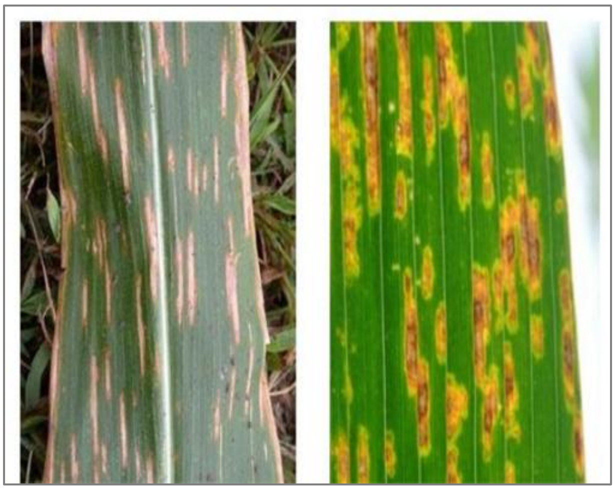

Fig 1: Gray leaf spot

Fig 2: Common rust

These lesions turn from pale gray to tan as the pathogens progress. Follow-up crop rotation with non-host crops is known as management. The paper's remaining portion is structured as follows: The related work of maize disease classification techniques is described in Sect. 2, several machine learning techniques are explained in detail in Sect. 3, the detailed procedure of the suggested method for maize leaf disease classification is presented in Sect. 4, the comparative analysis of all the classification techniques is presented in Sect. 5, and finally The study is concluded with the future scope of the work in Sect. 6.

Fig 3: Bright Leaf

Fig 4: Healthy leaf

MATERIALS AND METHODS

For disease classification, the Random Forest algorithm is chosen as it is a widely used machine learning method known for its high accuracy and ease of use. The dataset consists of four categories of maize leaf images: Gray leaf spot, Common rust, Blight, and Healthy. These images are collected and pre-processed to ensure consistency and quality, which includes resizing them to a uniform dimension (128×128), normalizing pixel values, and applying image augmentation techniques such as rotation and flipping to improve model robustness. After preprocessing, the dataset is divided into training and testing sets to evaluate the model's performance. Feature extraction techniques like texture analysis and color histograms are used to capture important visual characteristics from the images, and these features are then used to train the Random Forest model for accurate classification. As illustrated in Fig. 5, the overall block diagram represents the workflow from data input to classification output, highlighting the stages of preprocessing, feature extraction, model training, and prediction.

***Corresponding Author:** Shalini Hanok

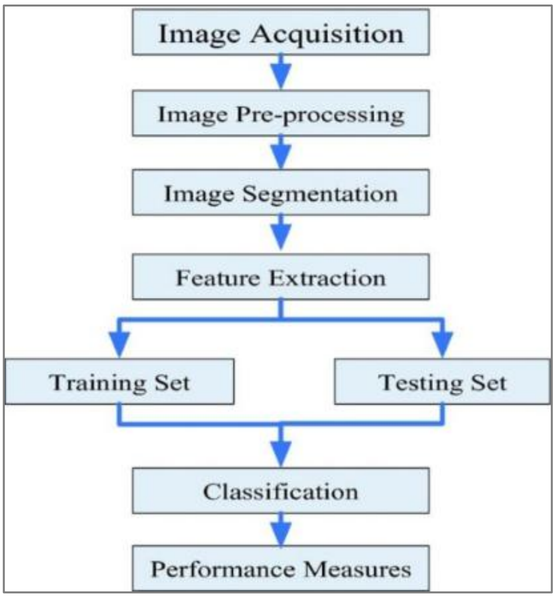

Fig 5: Block Diagram of the Model

Once the model is trained, its performance is evaluated using metrics such as accuracy, precision, recall, and F1-score to ensure reliability. To further improve the model's generalization capability, hyperparameter tuning is performed, adjusting parameters like the number of trees and maximum depth of the forest. Cross-validation techniques are also applied to confirm that the model performs well on unseen data. A user-friendly web application is then developed using the Python Flask framework, allowing users to upload maize leaf images for classification. The trained Random Forest model is integrated into the application, which processes the uploaded image and instantly predicts whether the leaf is healthy or affected by a specific disease. The application is deployed on a server to ensure accessibility across multiple devices, and it is thoroughly tested for performance and usability with different image inputs. When a user uploads an image, the system preprocesses it, extracts relevant features, and uses the trained Random Forest model to classify the leaf condition. The output is displayed to the user, indicating whether the leaf is healthy or diseased, along with the specific disease category if applicable. The system is also designed for continuous improvement, allowing updates such as incorporating more data or adding new disease categories to maintain accuracy over time. The dataset used consists of labeled maize leaf images, which serve as the foundation for both training and testing.During preprocessing, resizing, ormalization, and label encoding are applied to convert categorical labels into numerical form suitable for machine learning algorithms. The Random Forest model, an ensemble of decision trees, processes these inputs and makes predictions based on aggregated outputs. After training, the model is fine-tuned to achieve optimal performance and then saved in formats such as .pkl or .joblib, enabling efficient reuse without retraining, particularly during deployment in the web application.

RESULTS

The operating system library is used to read folders and determine file paths. The built-in NumPy library is used to convert data into array format, while the cv2 library helps in reading images. The pickle library is used to convert stored data into byte format, and the Scikit-learn (sklearn) library provides machine learning algorithms for model development and evaluation.

1. Accuracy:

The proportion of correct predictions to the total number of predictions is known as accuracy.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \quad ...1$$

2. Precision:

Precision is the ratio of true positives to the sum of true positives and false positives.

$$\text{Accuracy} = \frac{\text{True Positive}}{\text{True Positive} + \text{False positive}} \quad ...2$$

3.Recall:

Recall is the ratio of true positives to the sum of true positives and false negatives.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False positive}} \quad ...3$$

4.F1 Score:

F1 score is the harmonic mean of precision and recall.

$$\text{Recall} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$$

As shown in Fig. 6, these equations are used to evaluate the performance of the classification model.

Fig 6: Python Model for classifying Model

***Corresponding Author:** Shalini Hanok

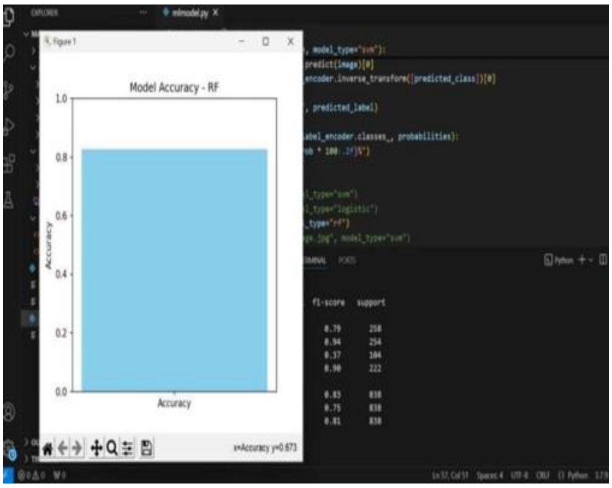

Fig 7: Plot for Accuracy

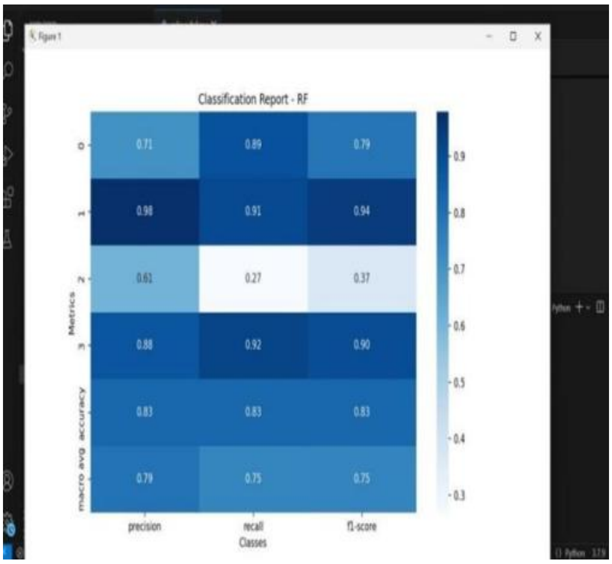

Fig 8: Confusion Matrix for prediction

DISCUSSION

The proposed maize leaf disease classification system demonstrates the effective application of machine learning techniques, particularly the Random Forest algorithm, for accurate and early detection of plant diseases. By utilizing a diverse dataset consisting of images categorized into Gray Leaf Spot, Common Rust, Blight, and Healthy classes, the system ensures reliable training and testing performance. Compared to traditional proaches and previously studied models such as KNN, Decision Tree, SVM, and Naive Bayes, the Random Forest model shows improved accuracy due to its ensemble nature, which combines multiple decision trees to reduce overfitting and enhance generalization. The achieved accuracy of 83% indicates a significant improvement over some existing works reported in the literature, which achieved around 79%. The preprocessing steps, including image resizing, normalization, augmentation, and feature extraction using texture and color analysis, contribute substantially to the robustness of the model. Furthermore, evaluation metrics such as accuracy, precision, recall, and F1-score provide a comprehensive assessment of the model's performance. The integration of the trained model into a Python Flask-based web application enhances its practical usability, enabling real-time disease prediction through image uploads. This user-friendly interface makes the system accessible to farmers and agricultural practitioners, allowing timely decision-making for disease management. Additionally, the system's scalability and ability to incorporate new data and disease categories make it adaptable to future agricultural challenges. Overall, the study highlights the potential of machine learning in precision agriculture, offering an efficient, cost-effective, and scalable solution for crop disease detection and management.

CONCLUSION

An efficient and reliable method for detecting and categorizing diseases in maize crops is achieved through the developed classification system using the Random Forest machine learning algorithm. This approach significantly reduces human error and enables timely disease management by automating the identification process, thereby eliminating the need for time-consuming manual inspection. As illustrated in Fig. 9, the comparison between different models highlights the improved performance of the proposed approach. The integration of a web-based interface using the Python Flask framework provides an accessible platform for farmers, agricultural experts, and researchers to upload leaf images and receive instant predictions, supporting better decision-making in crop management. The system demonstrates how artificial intelligence in agriculture can enhance productivity and minimize losses by combining machine learning techniques with modern web technologies. Furthermore, the ability to quickly identify healthy leaves and classify diseases such as Gray Leaf Spot, Common Rust, and Blight improves the efficiency of disease control. Its scalability and user-friendly design make it particularly beneficial for users in remote areas with limited access to expert guidance and advanced tools. Overall, the proposed system contributes to advancements in agricultural technology and establishes a foundation for future enhancements, including extending disease classification to other crops and incorporating treatment recommendations. By enabling early disease detection, the system plays a crucial role in improving agricultural practices, ensuring higher crop yields, and promoting sustainable farming. Additionally, when compared to existing literature where an accuracy of 79% was achieved, the proposed Random Forest model attains an improved accuracy of 83%, demonstrating its effectiveness.

***Corresponding Author:** Shalini Hanok

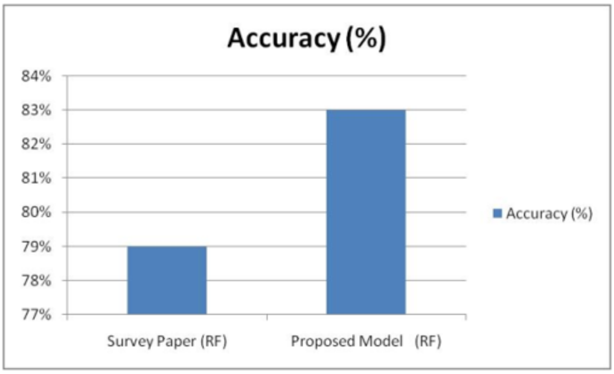

Fig 9: Comparison between the models

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to our project guide, **Ms. Shalini Hanok**, for her valuable guidance, continuous support, and encouragement throughout the completion of this project. Her insights and suggestions greatly contributed to the successful development of this work.

We also extend our appreciation to our team members, **Gagana S L, Ganavi B N, Akash L, and Mridula M Gowda**, for their cooperation, dedication, and collaborative efforts during all phases of the project. Their teamwork and commitment played a vital role in achieving our objectives.

We are grateful to the faculty members of the **Department of Electronics and Communication Engineering, ATME College of Engineering**, for providing the necessary resources, guidance, and a supportive academic environment.

REFERENCES

1. L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
2. A. Khan and D. S. Kim, “A machine learning-based maize disease classification using deep learning models,” *International Journal of Computer Applications*, vol. 175, no. 10, pp. 31–38, 2020, doi: 10.5120/ijca2020920447.
3. S. Mohan and S. Rajasree, “A review of machine learning in agriculture for plant disease detection,” *International Journal of Engineering and Technology*, vol. 7, no. 3, pp. 1310–1314, 2019, doi: 10.14419/ijet.v7i3.20695.
4. P. B. Pathak and V. Gupta, “Agricultural machine learning for disease prediction algorithms,” *International Journal of Computer Applications*, vol. 171, no. 8, pp. 20–25, 2022, doi: 10.5120/ijca2022923043.
5. T. Richey and M. Shrivalkar, “Machine learning techniques for plant disease detection,” in *Proc. IEEE Int. Conf. Machine Learning and Applications (ICMLA)*, 2019, pp. 328–333.
6. R. Sharma, A. Singh, K. Verma, N. Z. Janjhi, M. Masud, E. S. Hia, and S. Verma, “Plant disease diagnosis and image classification using deep learning,” *Computers, Materials & Continua*, vol. 71, no. 2, pp. 2127–2144, 2022, doi: 10.32604/cmc.2022.020017.
7. K. P. Panigrahi, H. Das, A. K. Sahoo, and S. C. Moharana, “Maize leaf disease detection and classification using machine learning algorithms,” in *Progress in Computing, Analytics and Networking*, Singapore: Springer, 2020, pp. 659–669.

***Corresponding Author:** Shalini Hanok