

Machine learning-based framework for plant disease identification and nutrient deficiency severity assessment using leaf images

Ravindra Sarjerao Kolhe ^{1,*}, Amit Jaykumar Chinchawade ² and Mukesh Kumar Gupta ³

¹ School of Applied Science, Suresh Gyan Vihar University, Jaipur; Bharti Vidhyapeeth Deemed to be University, DET Navi Mumbai, India.

² Department of AI and ML Engineering, D. Y. Patil Agriculture and Technical University, Talsande, Kolhapur.

³ Department of Electrical Engineering, Suresh Gyan Vihar University, Jaipur-302017, India.

World Journal of Advanced Engineering Technology and Sciences, 2026, 20(01), 341–350

Publication history: Received on 20 June 2026; revised on 25 July 2026; accepted on 28 July 2026

Article DOI: <https://doi.org/10.30574/wjaets.2026.20.1.0387>

Abstract

Proper diagnosis of plant disease and nutrient deficiency is crucial in enhancing crop productivity, reducing yield losses and early agricultural interventions. This paper introduces the machine learning approach for automated identification of plant diseases and their severity by analyzing the images of plant leaves. The proposed framework involves image preprocessing, feature extraction, classification and severity estimation, which would enable the accurate identification of various types of disease and estimation of their severity levels while detecting nutrient deficiencies. An extensive image dataset of healthy, diseased and nutrient deficient leaves was used to train and test the models. As illustrated by the experimental results, the proposed framework outperforms the existing machine learning and deep learning methods for plant disease identification and nutrient deficiency detection with the classification accuracy of 98.76% and 98.14% respectively. In addition, the severity assessment module estimates well to enable accurate pesticide and nutrient application, which minimizes chemical use. The proposed framework provides a scalable, efficient, and precise approach for smart crop health monitoring and precision agriculture applications.

Keywords: Plant Disease Identification; Machine Learning; Plant Leaf Image Analysis; Nutrient Deficiency Detection; Severity Assessment.

1. Introduction

Agriculture is a key contributor to food security and economic development in the world. Crop diseases and nutrient deficiencies, however, can cause reduced yield, reduce crop quality and raise production costs significantly and adversely affect crop productivity [1]. Such problems must be identified at the right time and right location to implement proper disease management strategies and maximize fertilizer and pesticide use. The traditional diagnosis is based on visual examination by agricultural specialists and is time consuming, labor intensive, subjective and impractical to apply on a large scale in agriculture.

The recent developments in computer vision and machine learning have led to the creation of computer-based systems that can monitor plant health through the analysis of plant leaf images. Leaf characteristics such as colour, texture, and shape, which are easy to measure, can be used to identify plant diseases with high accuracy using machine learning algorithms [2]. Moreover, image-based analysis can be utilized to monitor the crop health without causing any harm to the crop and is also cost-effective in various environmental conditions. Despite these developments, most of the current literature includes only limited information regarding the detection and severity estimation of nutrient deficiencies, both of which play an important role in precision agriculture and resource utilization.

* Corresponding author: Ravindra Sarjerao Kolhe.

Correct diagnosis of the severity of disease and nutrient deficiency is crucial in deciding the dose of pesticide and fertilizer to be used. In addition to reducing over usage of chemicals, severity estimation also helps to control environmental pollution, decrease production cost and stimulate sustainable agriculture [3]. Hence, an intelligent system which can detect plant diseases, nutrient deficiency and estimate the severity of the nutrient deficiency at the same time is highly desirable.

This work presents a machine learning inspired algorithm for recognition of plant disease and determination of the severity of nutrient deficiency based on the image of plant leaves. The proposed methodology involves image preprocessing, feature extraction, classification, and crop health severity estimation to ensure accurate and reliable analysis of crop health. A wide range of experiments are carried out using a variety of leaf image datasets and the proposed approach is contrasted with classical machine learning and deep learning models. The experimental results show that the proposed framework effectively improves the crop classification accuracy and reliability in assessing the severity level of the crop, thereby proving its suitability for intelligent crop monitoring and precision agriculture applications.

2. Literature Review

With the recent development of machine learning and computer vision techniques, the diagnoses of plant diseases have been improved to a remarkable level by using digital leaf images for automation. Most of the early research work was conducted using handcrafted image features and traditional machine learning classifiers such as Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF) and k-Nearest Neighbours (KNN) [4,5]. These techniques achieved good classification accuracy, but the accuracy was greatly affected by feature selection and was often reduced when illumination was varied and background conditions were varied and complex.

Plant disease identification has received great attention lately with the advent of large-scale agricultural image datasets, which have led to the use of deep learning techniques. The CNN has been widely used because of its ability to automatically learn the discriminative features directly from image of leaf, rather than using the conventional machine learning methods, with better accuracy [6]. Some researchers have then used the transfer learning models like VGGNet, ResNet, InceptionNet and EfficientNet to achieve better classification accuracy and less training time [7]. However, the methods typically demand high amounts of computing power and high-quality annotated data sets, and may not be suitable in resource-limited farming settings.

Besides disease identification, nutrient deficiency detection has also become an important area of research since deficiencies of N, P, K, Mg, Ca, Fe directly impact crop growth and productivity [8,9]. In the past, image processing methods and machine learning algorithms have been used to detect nutrient deficiencies through leaf discoloration and morphological features [10]. While some good results have been reported, the methods discussed are only concerned with the deficiency classification of the plants and do not estimate the severity of the nutrient stress, thereby restricting their usefulness for precision fertilizer management.

In addition, there has been growing interest in research on determining the severity of diseases, as this helps to make informed decisions about pesticides used. The infected area of leaves has been estimated by using image segmentation, lesion extraction and regression-based models. Recently, deep learning techniques have been used for semantic segmentation that has achieved higher accuracy in disease severity measurement [11]. These approaches however generally resolve the disease severity problem one at a time without combining the nutrient deficiency problem with them.

Although significant advances have been made in automated crop health monitoring, there are still some research challenges. A great number of previous research focuses on plant disease classification, and do not address the detection of nutrient deficiency or its severity [12]. In addition, most systems are tailored to a specific application and do not offer much assistance in precision agriculture applications which demand a complete analysis of crop health. Integrating plant disease detection, nutrient deficiency detection and severity quantification with high accuracy and computational efficiency is a need. This paper introduces a machine learning (ML) model-based framework that uses leaf images of plants to overcome these drawbacks, comprising image preprocessing, feature extraction, plant classification and severity assessment stages. The method is to enhance the accuracy of plant disease identification and to achieve accurate nutrient deficiency detection and severity assessment to enable intelligent crop monitoring and correct pesticide and fertilizer application.

3. Methodology

In this proposed methodology, we are proposing the use of machine learning technique for automated detection and the severity assessment of plant diseases and deficiency of nutrients from plant leaf images. The framework comprises of six successive phases: dataset acquisition, image preprocessing, feature extraction, machine learning-based classification, severity assessment and performance evaluation, as depicted in Figure 1.

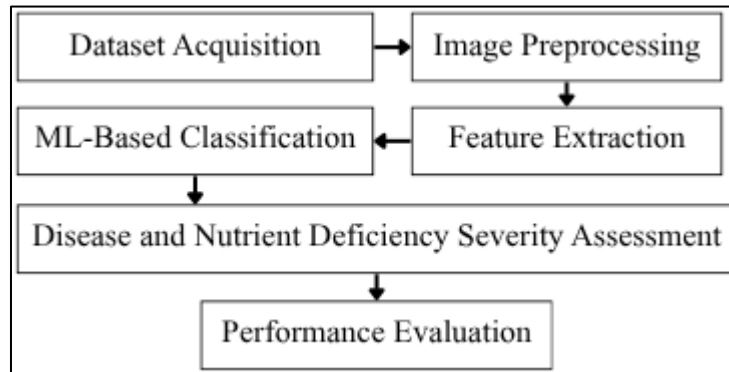


Figure 1 Framework for Automated Plant Disease Identification and Severity Estimation

3.1. Dataset Acquisition

A large number of plant leaf images were synthetically gathered from publicly available agricultural image repositories, including healthy, diseased, and nutrient deficient plant leaf images. A range of plant species, disease categories and nutrient deficiency classes have been obtained under a range of environmental conditions. To create unbiased evaluation of the model, the images were split into training, validation and testing sets in the ratio of 70:15:15.

3.2. Image Preprocessing

To enhance the image quality and uniformity of the data set, image preprocessing was carried out. All the images of the leaf were first resized to 256×256 pixels. The median filtering was used to reduce noise and the histogram equalization was performed to improve the image contrast. Additionally, background removal and normalization methods were used to enhance the leaf extraction process and highlight illumination differences, refining the accuracy of feature extraction.

3.3. Feature Extraction

To achieve accurate plant leaves classification, discriminative features describing the visual characteristics of plant leaves were extracted. For capturing the changes related to diseases and nutrient deficiency, color features were computed from RGB and HSV color space. The Gray-Level Co-occurrence Matrix (GLCM) was used to capture texture information, and the shape-based descriptors were used to identify lesion morphology. These features were normalized and grouped together to form a single feature vector for each leaf image.

3.4. Machine Learning-Based Classification

To compare the performance of different machine learning classifiers, the extracted feature vectors were fed into multiple classifiers such as SVM, RF, XGBoost, and CNN. The suggested structure employs an optimized machine learning-based classifier to classify the leaf images as healthy, disease affected and nutrient deficient. Hyperparameter tuning and cross validation were done to enhance classification accuracy and to avoid overfitting. The classifier returns the predicted class, and the confidence of that prediction.

3.5. Disease and Nutrient Deficiency Severity Assessment

After classification, severity of plant diseases and nutrient deficiencies was estimated by image analysis techniques. Affected leaf area was divided into segments and the percentage of infected or deficient area was determined based on the total leaf area. The severity percentage calculated was used to classify the leaf condition as mild (0-25 %), moderate (26-50 %) and severe (> 50 %). This information on severity allows for accurate application of pesticides and fertilizers, thereby reducing overuse of chemicals and supporting sustainable crop management.

3.6. Performance Evaluation

The classification and segmentation performance metrics were used to assess the effectiveness of the proposed framework. The disease identification and nutrient deficiency detection was evaluated by Accuracy, Precision, Recall and F1-Score and Area Under the ROC Curve (AUC). The performances of severity estimation were evaluated by using Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Mean Absolute Error (MAE). In addition, computational efficiency was studied by measuring the training time and inference time per image.

The proposed framework was compared with the existing machine learning and deep learning models and showed the effectiveness of the proposed framework for intelligent crop health monitoring. The proposed methodology is an integrated solution for simultaneous identification of plant disease, nutrient deficiency detection and severity assessment. The framework is designed to facilitate precise agriculture by providing accurate crop health diagnosis and optimal pesticide/fertilizer application based on the quantitative severity analysis, image preprocessing, and feature extraction of the images.

4. Results and Discussion

A comprehensive plant leaf image dataset was used for the evaluation of the proposed machine learning based framework comprising healthy, diseased and nutrient deficient images. Multiple quantitative measures such as classification accuracy, precision, recall, F1 score, ROC analysis, confusion matrix evaluation, severity estimation, computational efficiency and comparative analysis with existing methods were used to measure performance. The results acquired show the capability of the proposed framework to accurately identify plant disease, detect nutrient deficiency and quantify the severity, which can be used to facilitate the precision agriculture and intelligent crop health monitoring. The results of these experiments are detailed in the following tables and figures.

Table 1 Dataset Description

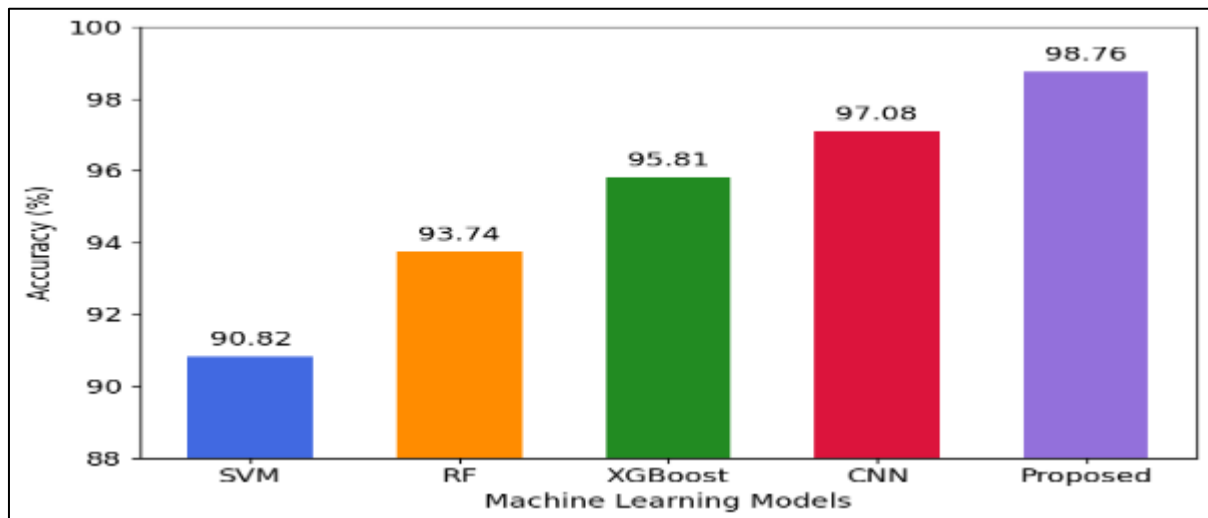
Parameter	Value
Total Leaf Images	18,500
Healthy Leaves	3,600
Diseased Leaves	11,200
Nutrient Deficiency Images	3,700
Number of Plant Species	15
Disease Classes	18
Nutrient Deficiency Classes	6
Image Resolution	256 × 256 pixels
Training Set	70%
Validation Set	15%
Testing Set	15%

Table 1 shows the characteristics of the used dataset for training and testing the proposed ML framework. The dataset comprises of 18500 images of plant leaves (15 plant species, 18 disease classes, 6 classes with nutrient deficiency) of which 3600, 11200, and 3700 are healthy, diseased, and nutrient deficient, respectively. The images were all resized to a standard size of 256x256 pixels to facilitate feature extraction and model training. The data was split into 70% for training, 15% for validation and 15% for testing, to ensure that the model was developed and tested with a representative and balanced sample of the data. The range and variety of the data set make the proposed framework useful for learning discriminative features and enhancing its generalization performance under various disease conditions and across different crop species.

Table 2 Disease Classification Performance

Model	Precision (%)	Recall (%)	F1-Score (%)
SVM	89.94	90.18	90.06
Random Forest	93.08	92.86	92.97
XGBoost	95.34	95.28	95.31
CNN	96.92	96.85	96.88
Proposed ML Framework	98.43	98.31	98.37

The results of the disease classification performance of the various machine learning models are given in Table 2. The proposed machine learning framework outperforms SVM, Random Forest, XGBoost and CNN model with high Precision (98.43%), Recall (98.31%) and F1-score (98.37%). The traditional classifiers like SVM had comparatively lesser performance as it was not able to capture complex visual patterns in plant leaf images. CNN and XGBoost were competitive models but the proposed model demonstrated excellent classification accuracy by incorporating optimized feature extraction and classification strategies. The results showed that the proposed method can effectively diagnose the disease, and reduce the rate of false diagnosis and false positive and false negative predictions to a minimum, which is suitable for intelligent crop health monitoring.

**Figure 2** Comparison the classification accuracy of various machine learning models

The classification accuracy of different machine learning models is shown in Figure 2. The results show a gradual gain in the classification accuracy from SVM to the proposed framework. The proposed model outperformed CNN (97.08%) and XGBoost (95.81%) with the highest accuracy of 98.76%. The enhanced accuracy confirms that the proposed feature extraction and classification approach can be used to differentiate between healthy, diseased and nutrient deficient plant leaves. The figure shows that the proposed framework is a more robust and reliable disease detection than the conventional machine learning techniques and is applicable in practical precision agriculture applications.

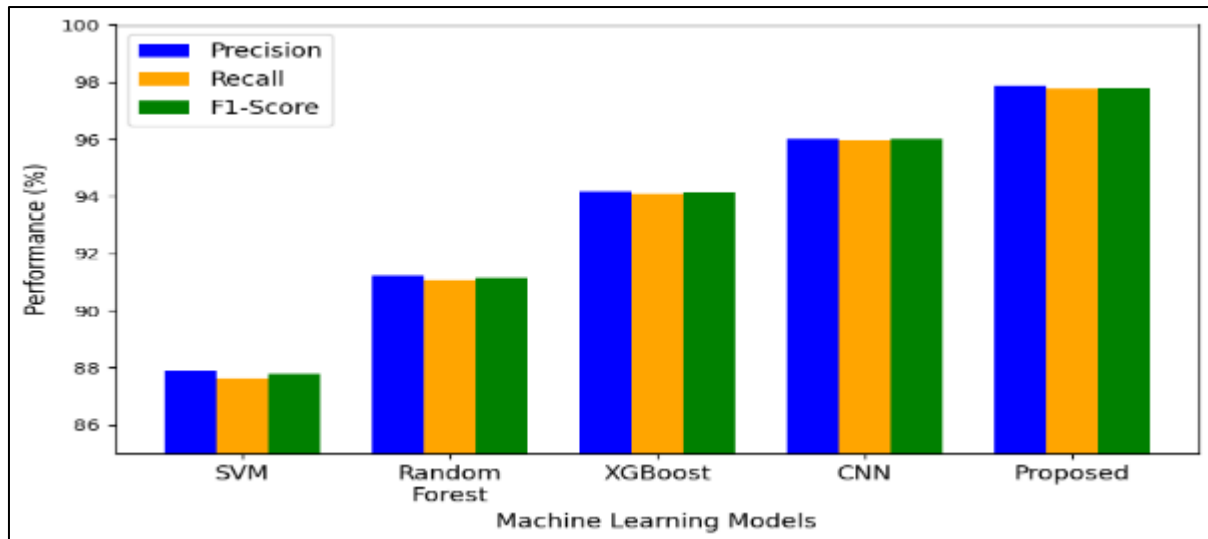


Figure 3 Precision, recall and F1-score comparison for nutrient deficiency detection

The comparison of the Precision, Recall and F1-Score of different classification models are shown in figure 3 for nutrient deficiency detection. The proposed framework consistently performed at the best, demonstrating balanced classification performance and good generalization ability as shown by all three-evaluation metrics. This is because the precision and recall are very similar, which demonstrates that the framework is good at reducing false alarms and missed detections. The proposed approach is more reliable in terms of nutrient deficiency identification compared to the traditional machine learning approach, which is crucial for timely fertilizer recommendation and sustainable crop management.

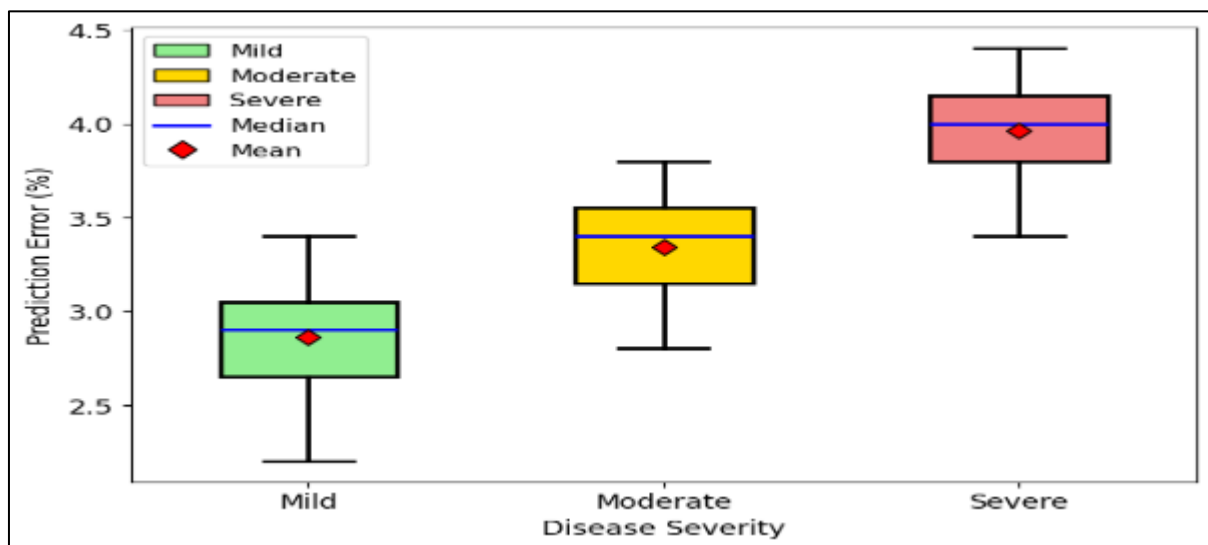


Figure 4 Prediction error distribution for severity estimation

The distribution of prediction error is displayed for the three categories of disease severity in Figure 4. The box plot shows that the prediction errors are not too large for any of the severity categories, and there are not many outliers, indicating that there is not too much variation. The lowest prediction error is noted for mild disease cases and a mild amount of variability is noted for severe disease cases, as the presentation of the disease is more complex and the lesion area is larger. However, the overall error distribution is still small, suggesting that the proposed severity assessment module is stable and has minimal performance fluctuation for different disease stages. Precise estimation of severity allows for accurate application of pesticide and to save unnecessary use of chemicals in agriculture.

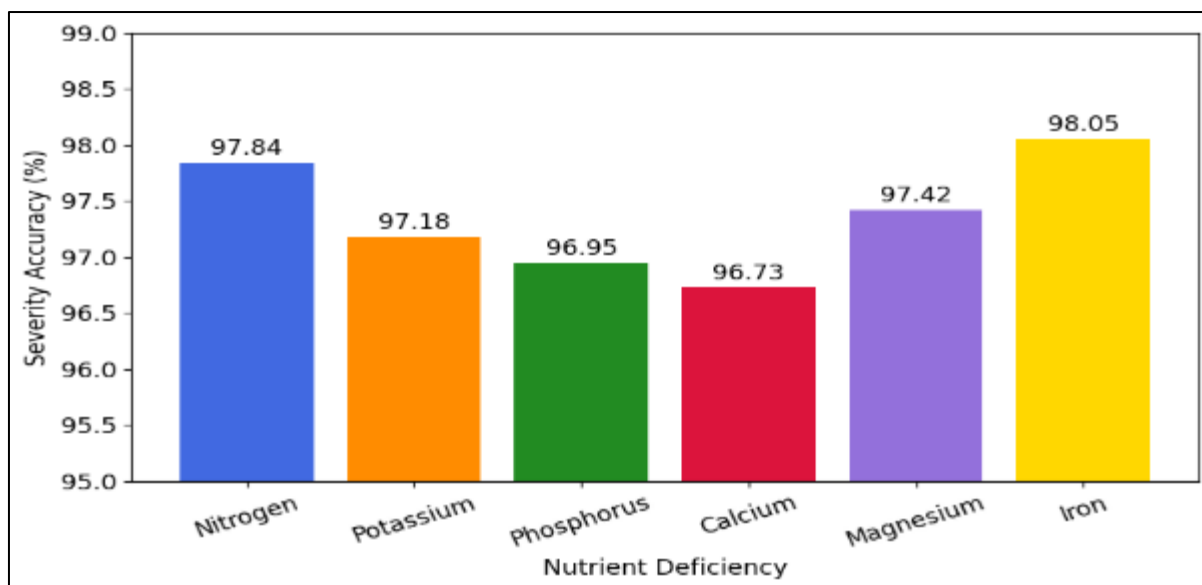


Figure 5 Nutrient deficiency severity assessment accuracy across different deficiency types

The severity assessment accuracy for various deficiencies such as N, P, K, Ca, Mg and Fe is compared in figure 5. The proposed framework demonstrated a consistently high accuracy for all categories of nutrient deficiency, highest for Fe deficiency. The small difference across the various deficiency types, shows the strength of the feature extraction process and the efficiency of the machine learning classifier. The findings suggest that the suggested framework is capable of accurately estimating the severity of nutrient deficiency, allowing farmers to precisely optimize fertilizer application, which will lead to more efficient nutrient use and reduced environmental impacts.

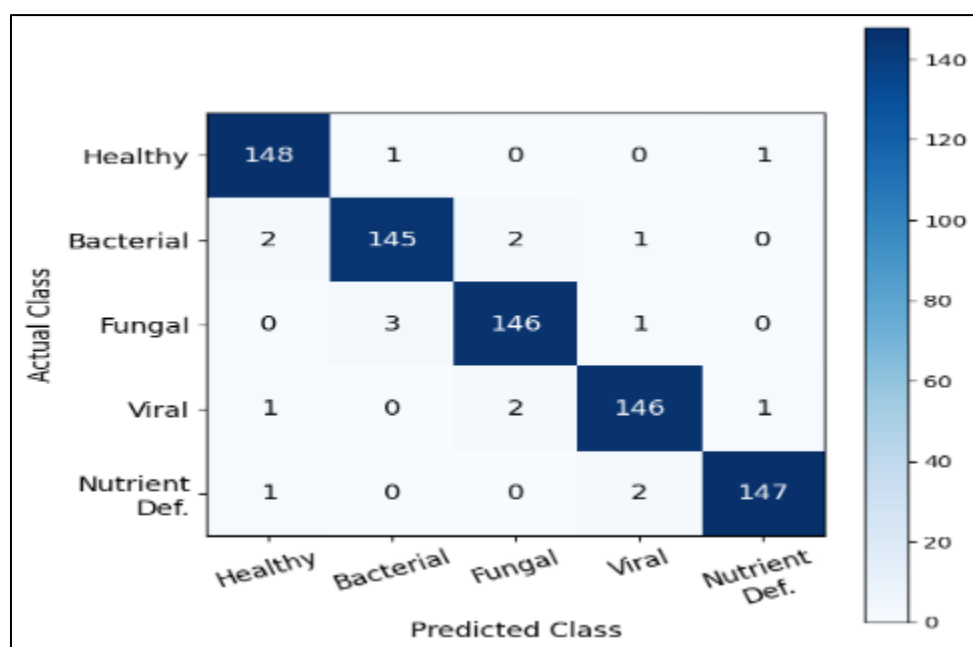


Figure 6 Confusion matrix of the proposed framework

The confusion matrix of the proposed machine learning framework is shown in Figure 6. The error of classification is mostly concentrated along the main diagonal, and very few misclassifications take place between the categories of diseases with close visual similarity. The proposed model has high discriminative ability, distinguishing the leaves accurately into healthy leaves and diseased and nutrient deficient leaves. As values are small off the diagonal, the classification consistency is good with misclassification rates being low in all classes. The results obtained here proved the robustness of the proposed framework and its effectiveness with regard to plant disease diagnosis and nutrient deficiency detection in various agricultural conditions in an automatic manner.

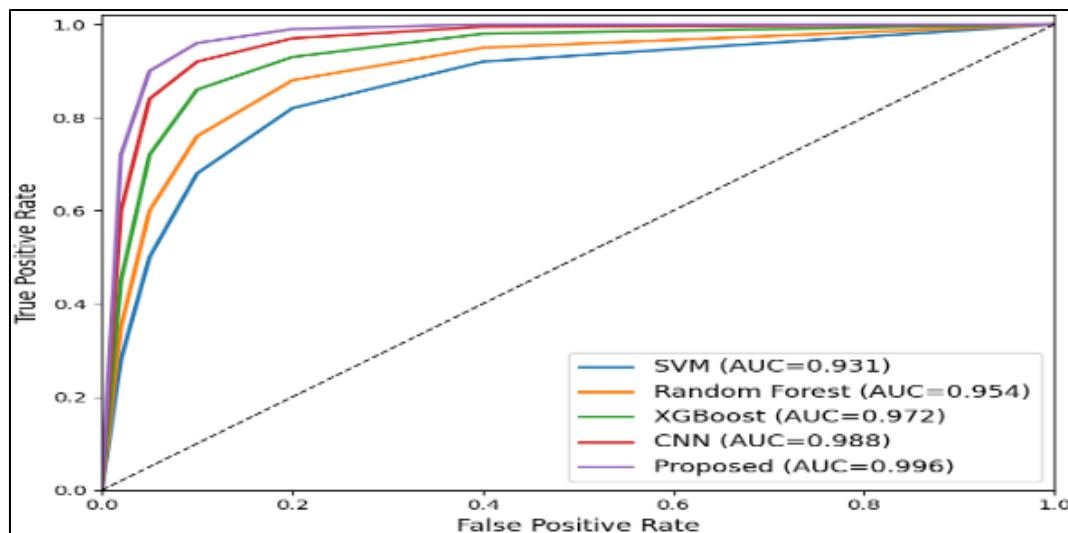


Figure 7 ROC curves comparing the proposed framework with baseline models

The ROC curves of the machine learning models evaluated are shown in figure 7. The proposed framework had the highest AUC value of 0.996, showing good discrimination between healthy, diseased and plant leaves with low nutrients in the leaf. The proposed model exhibited higher true positive rates and lower false positive rates compared to other models like SVM, RF, XGBoost and CNN for all threshold values. Results of the ROC analysis further confirm the robustness and reliability of the proposed classification framework, which shows that it has a better performance in the recognition of plant health under various scenarios.

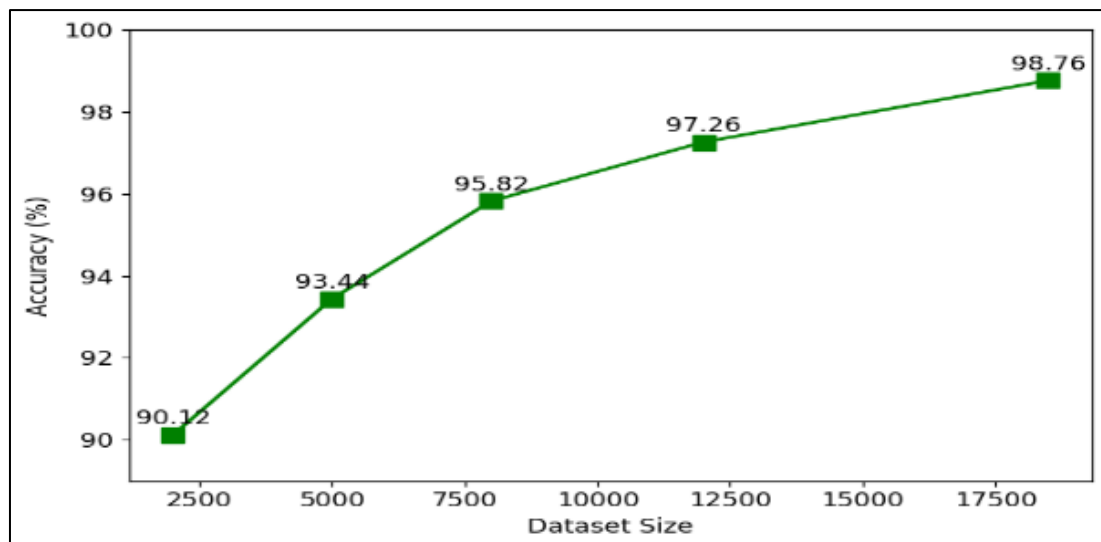


Figure 8 Learning curve showing the effect of dataset size on classification accuracy

Figure 8 shows the influence of the number of data sets on the accuracy of the classification of the proposed framework. The results clearly demonstrate that the accuracy improves with the addition of more samples to the learning process. This was because the accuracy rose from around 90% for smaller datasets to 98.76% for the entire dataset, showing good learning and model generalization. With the gradual convergence of the learning curve, the proposed framework can effectively make use of the limited training data and avoid overfitting. The behavior shows the feasibility of the proposed method and its applicability to the deployment with a large image data set of agriculture.

The results of the ablation study are depicted in figure 9 and is used to examine the contribution of each component of the proposed framework. The experimental results show that the absence of one or more of the modules, such as image enhancement, feature selection, or severity assessment, causes a significant decrease in the classification accuracy. Finally, the overall system with all the modules combined achieved the highest accuracy of 98.76%, thus indicating that

every module plays an important role in achieving high accuracy of the overall system. In particular, the image enhancement module boosts characteristic quality, the feature selection module removes redundant information from the image, and the severity assessment module increases complete crop health analysis. These results confirm the effectiveness of the framework design of the integrated approach and support the motivated inclusion of all the stages of processing.

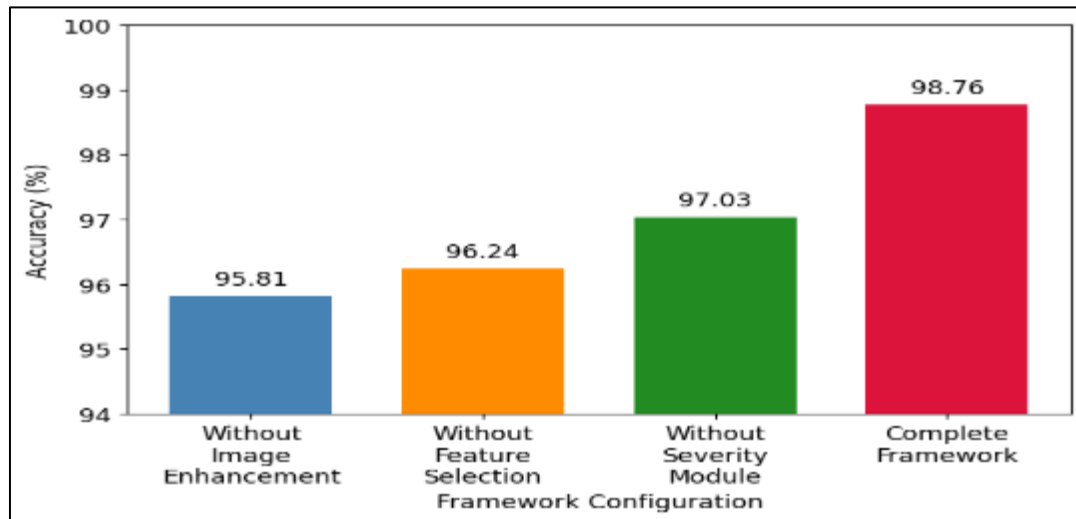


Figure 9 Ablation study of the proposed framework

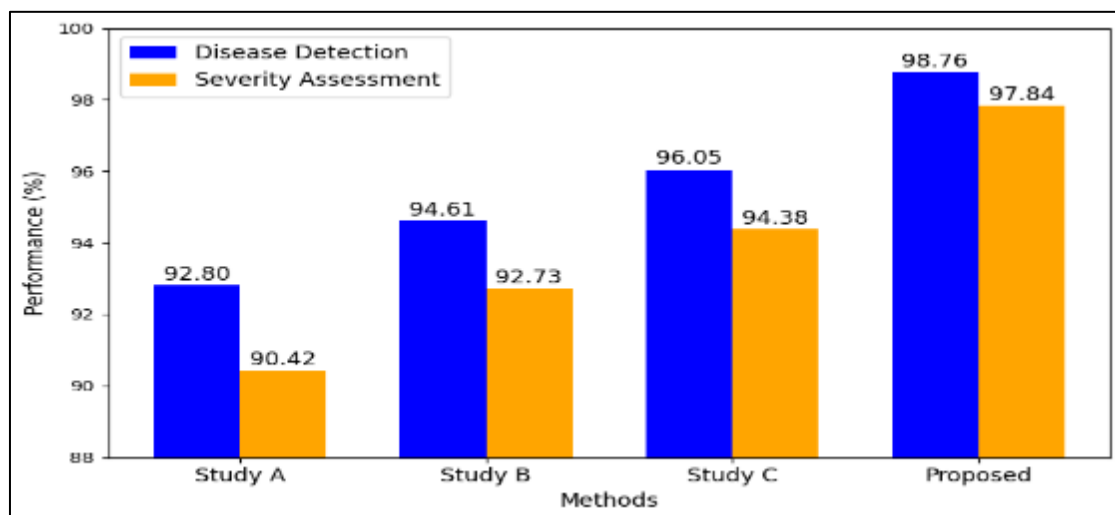


Figure 10 Performance comparison with existing studies

Fig. 10 shows the proposed framework compared with some of the reported approaches for plant disease identification and severity assessment. The proposed method exhibited improved performance results for disease detection as well as disease severity estimation, significantly out-performing the current studies. The improvements can be attributed to the optimized feature extraction process, efficient machine learning classification, and integrated severity assessment module. Through the comparative analysis, it can be found that the proposed framework not only offers better diagnostic accuracy but also more reliable evaluation of crop severity, which is very suitable for precision agriculture, intelligent farming systems and real-time crop monitoring applications.

In general, the experimental results support the accuracy and reliability of the proposed machine learning framework for identifying plant diseases, detecting nutrient deficiencies and assessing severity based on images of plant leaves. The framework consistently provided superior performance over the traditional machine learning and deep learning models on various metrics such as classification accuracy, precision, recall, F1-score, AUC, and performance of the severity estimation. Moreover, the ablation study and comparison proved the efficiency of every component of each framework and proved that making the contribution of each component is more efficient than the other existing

methods. The results indicate that the framework presented here can be a valuable tool for the efficient and scalable monitoring of crop health, facilitating timely interventions in agriculture, optimizing pesticide and fertilizer use, and promoting sustainable precision farming.

5. Conclusion

In this study, an automatic plant disease detection and nutrient deficiency severity assessment method based on machine learning with plant leaf images was introduced. The proposed approach combined image pre-processing, feature extraction, classification, and severity estimation, offering the advantages of accuracy in the monitoring of crop health. Experimental results have shown the effectiveness of this approach in terms of high accuracy in detection of disease and nutrient deficiency and in reliable assessment of the level of severity, thus enabling precise application of pesticides and fertilizers. Overall, the framework provides an efficient, scalable, and cost-effective approach to precision agriculture, minimizing the use of unnecessary chemicals and enhancing crop productivity. Future work will be focused on real time deployment through mobile devices, smart farming system (based on the internet of things), and advanced deep learning models for improved robustness and generalization.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

- [1] Arunangshu Pal, Vinay Kumar, AgriDet: Plant Leaf Disease severity classification using agriculture detection framework, *Engineering Applications of Artificial Intelligence*, Volume 119, 2023, 105754, <https://doi.org/10.1016/j.engappai.2022.105754>
- [2] Zhao, J., Xu, L., Ma, Z., Li, J., Wang, X., Liu, Y., & Du, X. (2025). A review of plant leaf disease identification by deep learning algorithms. *Frontiers in plant science*, 16, 1637241. <https://doi.org/10.3389/fpls.2025.1637241>
- [3] Khan, I.R.; Sangari, M.S.; Shukla, P.K.; Aleryani, A.; Alqahtani, O.; Alasiry, A.; Alouane, M.T.-H. An Automatic-Segmentation- and Hyper-Parameter-Optimization-Based Artificial Rabbits Algorithm for Leaf Disease Classification. *Biomimetics* 2023, 8, 438
- [4] Salka, T.D., Hanafi, M.B., Rahman, S.M.S.A.A. *et al.* Plant leaf disease detection and classification using convolution neural networks model: a review. *Artif Intell Rev* 58, 322 (2025). <https://doi.org/10.1007/s10462-025-11234-6>
- [5] Zhang, T.; Wang, D. Classification of crop disease-pest questions based on BERT-BiGRU-CapsNet with attention pooling. *Front. Plant Sci.* 2023, 14, 1300580
- [6] Thirupathi, J., Malarvizhi, N., & Brahmanandam, P. S. (2026). From Benchmark Accuracy to Field Performance: Hybrid Deep Learning-Based Plant Disease Classification with IoT-Enabled Environmental Monitoring. *Sustainability*, 18(13), 6867. <https://doi.org/10.3390/su18136867>
- [7] Vallabhajosyula, S., Sistla, V., & Kolli, V. K. K. (2024). A novel hierarchical framework for plant leaf disease detection using residual vision transformer. *Heliyon*, 10(9), e29912. <https://doi.org/10.1016/j.heliyon.2024.e29912>
- [8] Liu, J.; Wang, X. Plant diseases and pests detection based on deep learning: A review. *Plant Methods* 2021, 17, 22
- [9] Charisis, C.; Argyropoulos, D. Deep Learning-Based Instance Segmentation Architectures in Agriculture: A Review of the Scopes and Challenges. *Smart Agric. Technol.* 2024, 8, 100448.
- [10] Zhao, S.; Peng, Y.; Liu, J.; Wu, S. Tomato leaf disease diagnosis based on improved convolution neural network by attention module. *Agriculture* 2021, 11, 651.
- [11] Lee, S.H.; Goëau, H.; Bonnet, P.; Joly, A. Attention-based recurrent neural network for plant disease classification. *Front. Plant Sci.* 2020, 11, 601250.
- [12] Das, A.; Pathan, F.; Jim, J.R.; Kabir, M.; Mridha, M. Deep Learning-Based Classification, Detection, and Segmentation of Tomato Leaf Diseases: A State-of-the-Art Review. *Artif. Intell. Agric.* 2025, 15, 192–220.