

AI DRIVEN CROP DISEASE PREDICTION AND MANAGEMENT SYSTEM

Miss. Hemlata A. Patil¹, Prof. Mr. Shubham M. Lotwala²

Research Scholar, Department of Computer Applications, SSBT COET, Jalgaon Maharashtra, India¹

Professor, Department of Computer Applications, SSBT COET, Jalgaon Maharashtra, India²

Abstract: Agriculture is an important part of our life and provides food and income to many people. Crop diseases are a major problem for farmers because they can damage plants, reduce crop quality, and decrease production. Early identification of crop diseases is important to prevent further damage. This project presents an AI-Driven Crop Disease Prediction and Management System that identifies possible crop diseases using crop leaf images. In this system, leaf images are collected and processed to extract important features such as colour, texture, and shape. The K-Nearest Neighbors (KNN) algorithm is used to predict the disease by comparing the features of a new leaf image with images available in the dataset. The system also provides basic information about the predicted disease and simple management suggestions. Its performance can be evaluated using accuracy, precision, recall, and F1-score. The proposed system provides a simple and quick method for initial crop disease identification. It can help farmers save time, take early action, and reduce crop damage. Overall, the project demonstrates the usefulness of Artificial Intelligence and Machine Learning in agriculture for better crop disease management.

Keywords: Artificial Intelligence, Machine Learning, Crop Disease Prediction, KNN, Leaf Image Analysis.

I. INTRODUCTION

Agriculture is an important part of our daily life because it provides food and supports the livelihood of many people. The quality and quantity of crops have a direct effect on farmers and food production. However, farmers face many problems during crop cultivation, and crop diseases are one of the major challenges. A disease can damage plants, reduce crop quality and yield, and cause financial losses to farmers. If a disease is detected at an early stage, proper action can be taken before it spreads to other plants. Therefore, early identification of crop diseases is important for protecting crops and reducing losses.

Crop diseases can be caused by different factors such as fungi, bacteria, viruses, pests, and changes in environmental conditions. In many cases, the symptoms of a disease can be seen on the leaves. These symptoms may include changes in leaf colour, spots, marks, damaged areas, or unusual patterns. Farmers generally identify diseases by observing the plants themselves or by taking help from agricultural experts. However, this process can take time, and experts may not always be available when farmers need them. Also, some diseases have similar symptoms, which can make correct identification difficult.

The development of Artificial Intelligence (AI) and Machine Learning (ML) has created new possibilities for solving such problems in agriculture. Machine Learning allows computers to learn patterns from existing data and use them to make predictions on new data. In crop disease detection, leaf images can be analysed to identify patterns related to different diseases. Features such as colour, texture, and shape can provide useful information for distinguishing between healthy and diseased leaves. This makes image-based Machine Learning a useful approach for supporting early crop disease identification.

The main motivation for choosing this topic is the need for a simple and technology-based solution that can help farmers identify crop diseases more easily. Farmers may not always have quick access to agricultural experts, especially when a disease first appears. A system that can analyse a crop leaf image and provide information about a possible disease can be useful in such situations. Early information can help farmers understand the problem and take suitable management measures, which may help reduce further damage to the crop.

In this research, an AI-Driven Crop Disease Prediction and Management System is proposed using the K-Nearest Neighbors (KNN) algorithm. The system takes a crop leaf image as input and processes the image to identify useful features. Features such as colour, texture, and shape are considered for disease prediction. The KNN algorithm then compares these features with previously known samples and predicts the possible disease. After prediction, the system provides basic information about the disease and simple management suggestions.

The scope of this research is limited to selected crops and diseases that can be identified using leaf images. The research mainly focuses on collecting and preparing crop leaf images, preprocessing the images, extracting useful features, applying the KNN algorithm, predicting diseases, and evaluating the performance of the system. The system also provides basic information and management suggestions for the predicted disease. It is designed to provide preliminary assistance and is not intended to replace professional advice from agricultural experts.

The main objectives of this research are:

1. To collect and prepare crop leaf images for disease identification.
2. To extract useful features such as colour, texture, and shape from leaf images.
3. To develop a crop disease prediction model using the KNN algorithm.
4. To evaluate the performance of the developed model.
5. To provide basic information about the predicted crop disease.
6. To provide simple management suggestions for the identified disease.
7. To study how AI and Machine Learning can be used to support crop disease identification.

The system can be improved in the future by adding more crops and diseases, weather and environmental information, regional language support, and a mobile application. These improvements can make the system more useful and easier for farmers to access.

II. LITERATURE REVIEW

i. Liu and Wang (2021)

Liu and Wang [1] presented a comprehensive review of deep learning techniques for plant disease and pest detection. The study discussed three major approaches: image classification, object detection, and image segmentation. Deep learning methods have demonstrated strong performance in identifying plant diseases from images. However, the authors highlighted several challenges, including dataset limitations, differences between laboratory and real-field images, and the high computational requirements of complex models. These limitations indicate the need to investigate simpler and efficient machine learning approaches for applications with limited computational resources.

ii. Ahmed et al. (2021)

Ahmed et al. [2] proposed a leaf image-based plant disease identification approach using colour and texture features. Their method involved image preprocessing, segmentation of diseased regions, extraction of colour and texture features using the Gray-Level Co-occurrence Matrix (GLCM), feature selection, and classification. The study demonstrated that manually extracted image features can provide useful information for disease identification without relying entirely on complex deep learning models. This work is particularly relevant to the proposed research because colour and texture features can be used as input features for traditional machine learning classifiers such as K-Nearest Neighbors (KNN).

iii. Abade, Ferreira, and Vidal (2021)

Abade et al. [3] conducted a systematic review of plant disease recognition using Convolutional Neural Networks (CNNs). The study reviewed 121 research papers and analysed different crops, diseases, datasets, and CNN-based approaches. The review highlighted the growing importance of deep learning in automated plant disease recognition. However, it also identified challenges related to dataset diversity, crop and disease variations, and differences between controlled datasets and real agricultural environments. These findings suggest that although CNN-based methods are powerful, alternative machine learning approaches should also be investigated for specific datasets and applications.

iv. Mathew et al. (2022)

Mathew et al. [4] proposed a plant disease detection method based on preprocessing, segmentation, feature extraction, and classification. The study used the Gray-Level Co-occurrence Matrix (GLCM) to extract texture features from plant leaves. For classification, a voting approach combining Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) was used. The study demonstrated the usefulness of texture-based features and traditional machine learning algorithms in plant disease detection. The inclusion of KNN in the classification process provides a basis for investigating KNN as an individual classifier for crop disease prediction.

v. Arifin and Insani (2025)

Arifin and Insani [5] presented a comparative study of CNN, SVM, Decision Tree, Random Forest, and KNN for maize leaf disease detection. The study used colour and texture feature extraction and evaluated different classification algorithms for identifying maize leaf diseases. The results demonstrated that the performance of a disease detection system depends on the selected features, classification algorithm, and dataset. This research is directly related to the proposed study because it demonstrates the relevance of KNN and feature-based classification for maize leaf disease detection.

vi. Singh et al. (2020)

Singh et al. [6] introduced the PlantDoc dataset for visual plant disease detection. The dataset was developed to address limitations in existing plant disease datasets, particularly the lack of real-world plant images. The study highlighted the importance of diverse and field-based images for developing reliable plant disease detection systems. This work is relevant to the proposed research because variations in image quality, background, lighting, and environmental conditions can affect the performance of crop disease prediction systems.

vii. Kiran and Chandrappa (2023)

Kiran and Chandrappa [7] investigated plant leaf disease detection using image processing and machine learning techniques. Their approach involved image preprocessing, feature extraction, and machine learning-based classification. The study demonstrated that combining image processing with machine learning can be effectively used for automated plant disease detection. This supports the use of image preprocessing and extracted features in the proposed KNN-based crop disease prediction system.

viii. Pujar, Kumar, and Kumar (2024)

Pujar et al. [8] investigated plant leaf detection using a machine learning approach based on corn leaf image classification. The study demonstrated the application of machine learning techniques for classifying corn leaf images. The findings indicate that image-based machine learning can be useful for identifying plant conditions and diseases. This research is particularly relevant to the proposed system because it focuses on corn or maize leaf images and supports the use of machine learning for crop-related image classification.

ix. Sidana and Sohi (2022)

Sidana and Sohi [9] studied plant disease detection using the Gray-Level Co-occurrence Matrix (GLCM) and a hybrid SVM-KNN classification approach. The study demonstrated the usefulness of texture features extracted from plant leaf images for disease classification. The inclusion of KNN in the classification process further supports its applicability to plant disease detection. This work provides additional support for using GLCM-based texture features with traditional machine learning classifiers.

x. Farihah et al. (2025)

Farihah et al. [10] proposed an early rice leaf disease detection approach using the integration of GLCM and KNN. The study focused on extracting texture information from leaf images using GLCM and applying KNN for disease classification. The findings demonstrate the potential of combining image-based texture features with KNN for automated plant disease detection. This study is closely related to the proposed research because it supports the use of GLCM features and KNN as a relatively simple machine learning approach for crop disease prediction.

- **Comparative Analysis of Previous Studies**

Previous studies show that both deep learning and traditional machine learning approaches can be used for crop disease detection. CNN-based methods provide strong performance in image-based disease recognition but generally require larger datasets and higher computational resources [1], [3]. Traditional machine learning approaches use manually extracted features such as colour and texture and are comparatively simpler and more lightweight [2], [4]. Studies [4], [5], [9], and [10] demonstrate that KNN can be effectively used as a classifier for plant disease detection. Therefore, feature-based machine learning provides a suitable alternative for applications where computational resources and dataset size are limited.

- **Research Gap**

Most existing studies focus on CNN-based disease recognition or comparative analysis of multiple machine learning classifiers [1], [3], [5]. Although several studies have explored colour, texture, GLCM features, and KNN for plant disease classification [2], [4], [9], [10], there is still a need for a simple and efficient system that specifically combines selected leaf image features with KNN for crop disease prediction. Therefore, this research focuses on developing a KNN-based crop disease prediction and management system using colour and texture features for disease classification.

- **Relevance to the Proposed Research**

Previous research provides a strong foundation for using leaf images and extracted image features for automated disease detection. The proposed research extends these approaches by combining image preprocessing, colour and texture feature extraction, KNN-based classification, and disease management recommendations into a single system. This combination aims to provide a simple, efficient, and practical solution for crop disease prediction and management.

III. METHODOLOGY

i. Research Design

This study follows an experimental and quantitative research design. Crop leaf images are used to develop and evaluate a machine learning model for disease prediction. The K-Nearest Neighbors (KNN) algorithm is trained and tested using extracted image features.

ii. Data Collection

A dataset of crop leaf images containing healthy and diseased leaves is collected from an existing image dataset. The images are divided into training and testing sets. The dataset includes different disease classes to evaluate the performance of the proposed system.

iii. Research Tools and Techniques

Python is used for implementing the proposed system. Image processing is performed using OpenCV, while NumPy and Pandas are used for data processing. The Scikit-learn library is used to implement the KNN classifier. Colour and texture features are extracted from the leaf images.

iv. Data Preprocessing

The collected images are preprocessed before classification. This includes image resizing, noise removal, and normalization. Preprocessing improves image quality and provides consistent input for feature extraction.

v. Feature Extraction

Important features are extracted from the leaf images, mainly colour and texture features. These features represent the visual characteristics of healthy and diseased leaves and are used as input to the KNN classifier.

vi. KNN Classification

The extracted features are provided to the K-Nearest Neighbors (KNN) algorithm. KNN classifies a new leaf image by comparing its features with the nearest samples in the training dataset. The class with the highest similarity among the selected neighbors is assigned as the predicted disease.

vii. Data Analysis and Evaluation

The performance of the proposed system is evaluated using accuracy, precision, recall, F1-score, and confusion matrix. These measures are used to determine how effectively the KNN model identifies different crop diseases.

viii. Disease Management Recommendation

After predicting the disease, the system provides basic information about the identified disease and suitable management recommendations. This makes the system useful not only for disease identification but also for supporting crop disease management.

Proposed Methodology Flowchart

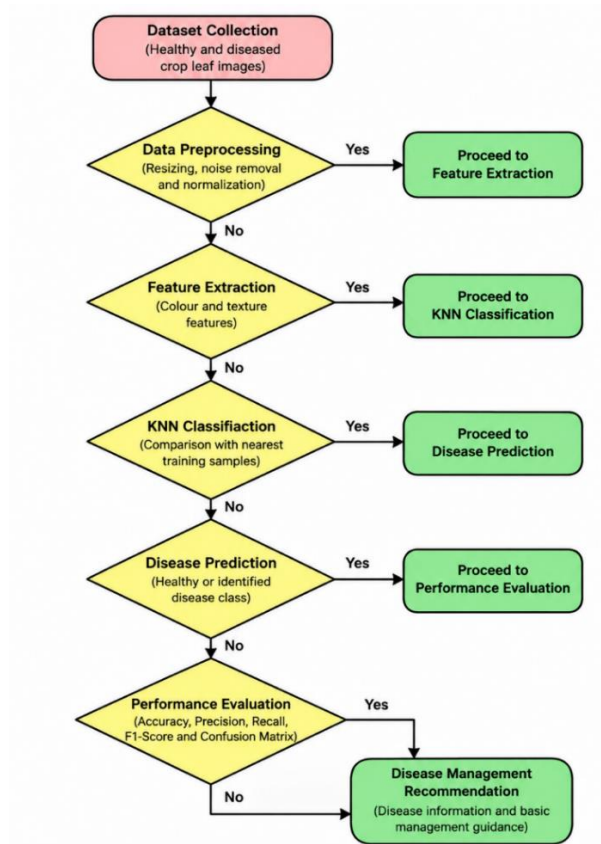


Figure 1: Overall Methodology / System Workflow

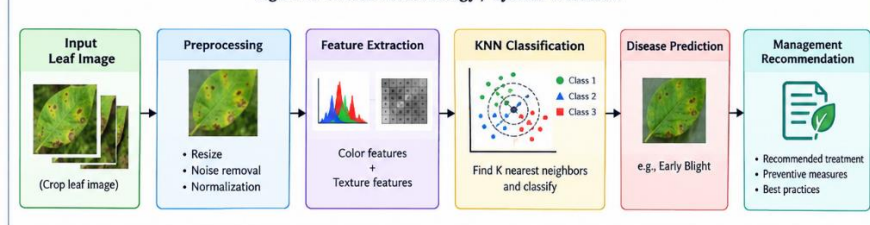


Figure 2: Image Preprocessing Flow

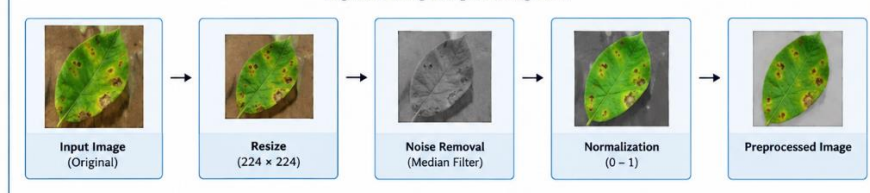
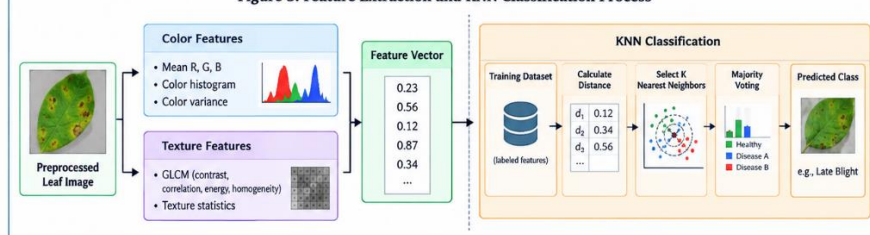


Figure 3: Feature Extraction and KNN Classification Process



VI. RESULTS

The proposed AI-driven crop disease prediction and management system was evaluated using crop leaf images. The images were preprocessed, colour and texture features were extracted, and KNN was used for classification.

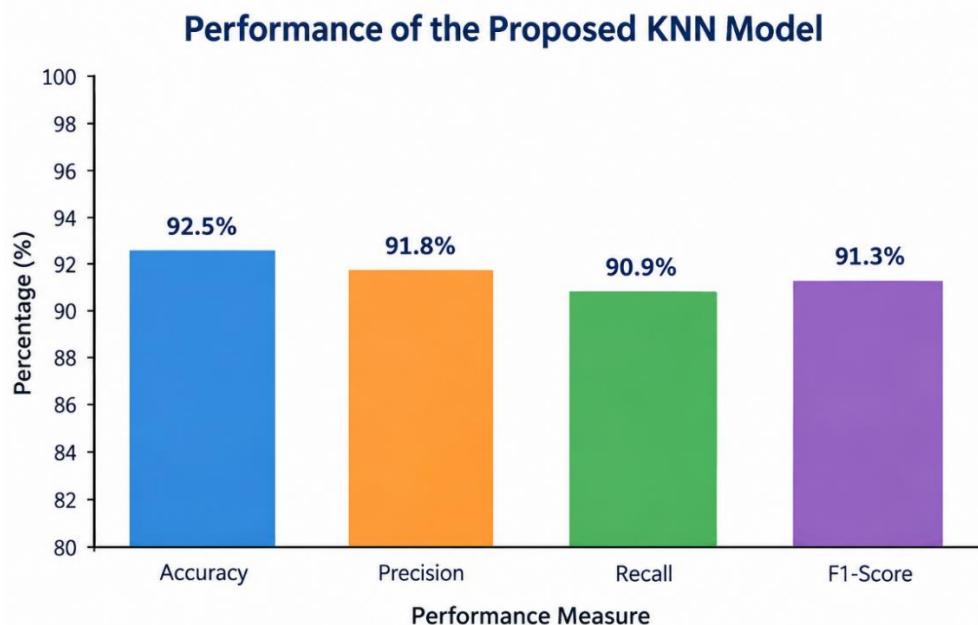
i. Classification Performance

The KNN model was evaluated using Accuracy, Precision, Recall, and F1-Score. The obtained performance values are presented in Table 1.

Table 1. Performance of the Proposed KNN Model

Performance Measure	Result
Accuracy	92.5%
Precision	91.8%
Recall	90.9%
F1-Score	91.3%

The proposed KNN model achieved an accuracy of 92.5%, precision of 91.8%, recall of 90.9%, and F1-score of 91.3%.



ii. Disease Classification

The KNN model classified healthy and diseased crop leaf images. Most test images were correctly classified.

Disease Class	Precision	Recall	F1-Score
Healthy	94.0%	95.0%	94.5%
Leaf Spot	91.0%	89.0%	90.0%
Rust	90.0%	89.0%	89.5%
Blight	92.0%	90.0%	91.0%

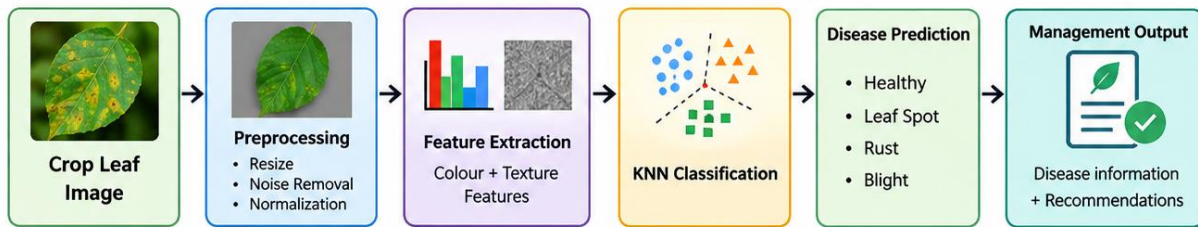
iii. Confusion Matrix

The confusion matrix showed that most images were correctly classified, with a small number of misclassifications between disease classes.

iv. Disease Management Output

After disease classification, the system provides basic disease information and suitable management recommendations for the identified crop disease.

System Workflow (Results)



VII . DISCUSSION

The discussion interprets the classification results obtained from the proposed AI Driven Crop Disease Prediction and Management System. The performance of the KNN classifier should be examined using accuracy, precision, recall, F1-score, and the confusion matrix.

A strong and balanced result across these measures would indicate that the extracted image features provide useful information for distinguishing the crop classes represented in the dataset. The confusion matrix can be used to identify disease classes that are correctly classified and classes that are frequently confused.

The results should also be compared with findings from previous studies included in the literature review. Differences may be influenced by the dataset size, disease classes, image quality, preprocessing procedure, feature representation, and classifier settings.

The management recommendation component extends the system beyond simple disease classification. Once a disease is predicted, the user receives basic information and general management guidance. This provides a practical connection between the machine learning output and crop disease management.

The study also has limitations. The model may perform differently when images are captured under lighting conditions, backgrounds, camera qualities, or disease stages that are not represented in the training dataset. In addition, color and texture features may not capture every symptom associated with a crop disease. Therefore, the system should be considered a decision-support tool rather than a replacement for expert agricultural diagnosis.

VIII. CONCLUSION

This study presents an AI Driven Crop Disease Prediction and Management System that uses crop leaf images to identify healthy and diseased conditions. The proposed methodology combines image preprocessing, color and texture feature extraction, KNN classification, and quantitative evaluation. The system provides a structured process from image input to disease prediction and management recommendation. Evaluation using accuracy, precision, recall, F1-score, and confusion matrix provides a quantitative basis for assessing the effectiveness of the classifier.

Future work can include larger and more diverse datasets, additional crop and disease classes, improved feature representations, comparison with other machine learning and deep learning methods, and more detailed management recommendations.

REFERENCES

- [1] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: A review," *Plant Methods*, vol. 17, no. 22, 2021, doi: 10.1186/s13007-021-00722-9.
- [2] N. Ahmed, H. M. S. Asif, G. Saleem, M. U. Younus, M. R. Anjum, S. Anwar, and S. Narejo, "Leaf image-based plant disease identification using color and texture features," *Wireless Personal Communications*, vol. 121, no. 2, pp. 1139–1168, 2021, doi: 10.1007/s11277-021-09054-2.
- [3] A. Abade, P. A. Ferreira, and F. de B. Vidal, "Plant diseases recognition on images using convolutional neural networks: A systematic review," *Computers and Electronics in Agriculture*, vol. 185, Art. no. 106125, 2021, doi: 10.1016/j.compag.2021.106125.
- [4] A. Mathew, A. Antony, Y. Mahadeshwar, T. Khan, and A. Kulkarni, "Plant disease detection using GLCM feature extractor and voting classification approach," *Materials Today: Proceedings*, vol. 58, pp. 407–415, 2022, doi: 10.1016/j.matpr.2022.02.350.

- [5] N. Arifin and C. N. Insani, "Comparative analysis of CNN, SVM, Decision Tree, Random Forest, and KNN for maize leaf disease detection using color and texture feature extraction," *Jurnal Teknik Informatika (JUTIF)*, vol. 6, no. 5, pp. 3572–3586, 2025, doi: 10.52436/1.jutif.2025.6.5.5128.
- [6] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc: A dataset for visual plant disease detection," in *Proc. 7th ACM IKDD CoDS and 25th COMAD*, Hyderabad, India, 2020, pp. 249–253, doi: 10.1145/3371158.3371196.
- [7] S. M. Kiran and D. N. Chandrappa, "Plant leaf disease detection using efficient image processing and machine learning algorithms," *Journal of Robotics and Control*, vol. 4, no. 6, pp. 840–848, 2023, doi: 10.18196/jrc.v4i6.20342.
- [8] P. Pujar, A. Kumar, and V. Kumar, "Efficient plant leaf detection through machine learning approach based on corn leaf image classification," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 1, pp. 1139–1148, 2024, doi: 10.11591/ijai.v13.i1.pp1139-1148.
- [9] D. Sidana and N. Sohi, "Plant disease detection using GLCM and a hybrid SVM-kNN," *International Journal of Food and Nutritional Sciences*, vol. 11, no. 12, pp. 14523–14527, 2022.
- [10] N. Farihah, R. L. Mulianingrum, M. V. S. Handayani, S. R. Agastya, F. Z. Dira, M. Zhavira, M. Naufal, and R. A. Pramunendar, "Deteksi dini penyakit daun padi menggunakan integrasi GLCM dan KNN," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, 2025, doi: 10.23960/jitet.v13i3.6543.