International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

# Leaf Disease Classification for Vegetable Farming using Artificial Intelligence

**Onyeachonam Dominic - Mario Chiadika[1], Ujerekre Ekoko[2]**

[1]Department of Electronic and Computer Engineering, Brunel University, London.
Email: *mario.chiadika[at]brunel.ac.uk*

[2]Department Computer Engineering, Delta State Polytechnic, Ogwashi - uku, Delta State
Email: *Uje.Ekoko[at]dspg.edu*

**Abstract:** ***This research paper explores the application of artificial intelligence (AI) techniques for the classification of leaf diseases in vegetable farming, utilizing Python as the primary programming language. The study aims to develop an accurate and efficient model to assist farmers in early detection and management of plant diseases, thereby improving crop yield and quality. By leveraging convolutional neural networks (CNNs), a dataset of diseased and healthy leaf images is processed to train and validate the AI model. The results demonstrate significant improvements in disease identification accuracy compared to traditional methods. The implementation of this AI - driven approach has the potential to revolutionize agricultural practices by providing a scalable, cost - effective solution for real - time disease monitoring and intervention.***

**Keywords:** Artificial intelligence, Leaf diseases, Vegetable farming, Convolutional neural networks, Real - time monitoring

## 1. Introduction

Leaf diseases impact many different crops, including wheat, rice, maize, soybeans, and vegetables, and they represent serious risks to global agricultural productivity. The primary causes behind these illnesses are a variety of pathogens, including bacteria, viruses, fungi, and phytoplasmas. The signs of these diseases include leaf spots, blights, rusts, mildew, and other conditions that lower photosynthetic activity, diminish yields, and degrade the quality of agricultural goods [1].

However, if leaf illnesses are discovered early on, farmers can stop disease outbreaks before they do serious harm to their crops by identifying and controlling the disease. Farmers can identify early signs of disease onset, such as leaf discoloration, lesions, and deformities, by putting proactive disease monitoring techniques into practice, such as routine field inspections and surveillance [2]. Then, early intervention strategies can be used to slow the spread of the disease and reduce yield losses. These strategies include resistant crop varieties, cultural techniques, and targeted fungicide applications.

Accurate classification of leaf diseases is essential for implementing targeted and effective disease management strategies. By correctly identifying the causal agents of leaf diseases, such as fungi, bacteria, viruses, or abiotic stressors, farmers can tailor treatment options to specific pathogens or environmental conditions. For example, certain fungicides may be effective against fungal pathogens but ineffective against bacterial diseases, highlighting the importance of accurate disease classification for selecting appropriate control measures [3]. Additionally, the classification of disease - resistant crop varieties enables farmers to proactively mitigate disease risks through crop selection and breeding programs.

Economic Benefits and Sustainability: Early leaf disease detection and classification offer significant economic benefits by reducing the costs associated with disease management and crop losses. By minimizing yield losses resulting from disease outbreaks, farmers can preserve their agricultural investments and maintain profitability. Moreover, by reducing the reliance on chemical inputs, such as pesticides and fungicides, early disease detection and classification contribute to sustainable farming practices and environmental stewardship. Sustainable disease management strategies, such as integrated pest management (IPM) approaches, emphasize the use of cultural, biological, and mechanical control methods in conjunction with chemical interventions, thereby minimizing environmental impacts while preserving crop health

The identification and classification of leaf diseases have been completely transformed by technological advancements such as artificial intelligence (AI), unmanned aerial vehicles (UAVs), and remote sensing [4]. Hyperspectral imaging and multispectral analysis are two examples of remote sensing techniques that make it possible to identify minute variations in leaf morphology and reflectance that are linked to disease symptoms. Convolutional neural networks (CNNs), support vector machines (SVMs), and deep learning models are examples of machine learning methods that provide automatic disease classification based on image - based attributes taken from digital leaf images. Farmers now have access to real - time, data - driven insights for proactive disease control and decision - making thanks to these technological advancements. The goal of this paper is to construct a leaf disease classification system for maize fields.

## 2. Background

The classification of leaf diseases has seen a rise in the use of computer vision algorithms by researchers in recent years, with the goals of enhancing crop output, optimizing disease management techniques, and improving disease detection accuracy. This section of this work presents a summary of earlier research on the use of computer vision techniques for the categorization of leaf diseases, emphasizing important approaches, conclusions, and problems that have been encountered in the field.

In order to distinguish between healthy and diseased leaves, early research on leaf disease classification using computer vision techniques concentrated on creating image processing algorithms and feature extraction approaches [5]. Leaf pictures were preprocessed and pertinent features, like texture, color, and shape characteristics, were extracted using conventional image processing techniques like thresholding, edge detection, and morphological procedures. These attributes were then used to train machine learning algorithms, such as k - nearest neighbors (KNN), decision trees, and support vector machines (SVM), to classify leaves into categories of healthy and unhealthy.

However, Research on leaf disease categorization has been transformed in recent years by the introduction of deep learning methods, especially convolutional neural networks (CNNs). Handcrafted feature extraction is no longer necessary because CNNs can automatically learn hierarchical features straight from raw pixel data. When it comes to classifying leaf diseases, numerous studies have shown that CNNs outperform more conventional machine learning techniques. Classification accuracy and generalization abilities have been further enhanced by the use of transfer learning, a method that makes use of pre - trained CNN models and refines them on particular leaf disease datasets [6].

## 3. Dataset Creation and Benchmarking

Research on computer vision - based leaf disease classification has advanced significantly thanks to the availability of high - quality leaf disease datasets. Large - scale datasets with hundreds of photos of both healthy and diseased leaves from various crop species have been collected and interpreted by researchers [7] [8]. These datasets are used as benchmarks to compare the efficacy of various feature extraction methods and assess how well certain classification algorithms perform. Open - access databases like PlantVillage and the Plant Disease Detection Dataset have made it easier for academics to collaborate and share knowledge.

Although computer vision techniques have demonstrated significant potential in automating the categorization of leaf diseases, there are still several obstacles and restrictions with current methods. Comprehending and resolving these obstacles is essential to enhance the efficiency and suitability of computer vision in agricultural settings. The main obstacles and restrictions that current methods of classifying leaf diseases using computer vision techniques encounter are covered in this paper.

- **Variability in Leaf Appearance:**
The intrinsic heterogeneity in leaf appearance caused by elements including lighting conditions, leaf orientation, and disease progression stages is one of the main obstacles to classifying leaf diseases. The symptoms and presentations of diseases in leaves can vary greatly, which makes it difficult to create reliable classification models that can reliably distinguish between healthy and diseased leaves in a variety of scenarios.

- **Limited Availability of Labeled Data:**
The scarcity of excellent labeled datasets for classification model training and validation represents a major obstacle as well. It can take a lot of time, money, and effort to gather and annotate large - scale databases of leaf photos with precise disease classifications. In addition, the variety of plant species, illnesses, and environmental factors makes gathering datasets more difficult and increases the risk of biases and problems with generalization in models that have already been trained.

- **Overfitting and Generalization:**
Machine learning - based leaf disease classification frequently faces the problem of overfitting, in which a classification model learns to memorize training data instead of generalizing patterns. When applied to unseen test data, overfitted models may perform well on training data but poorly in real - world scenarios due to their inability to generalize. Reliable disease detection and classification depend on the model's resilience and generalization abilities across various crop species, disease kinds, and environmental variables.

- **Interpretability and Explainability:**
Another major problem in leaf disease classification is interpreting and explaining the decisions made by machine learning models, especially in complicated deep learning architectures like convolutional neural networks (CNNs). Because these models function as "black boxes, " it is challenging to comprehend the underlying characteristics and decision - making procedures that determine classification results. Interpretability issues could prevent end users, such farmers and other agricultural stakeholders, from adopting and believing computer vision - based disease classification systems to be reliable [9] [10].

- **Real - Time Processing and Deployment:**
The implementation of computer vision - based systems for classifying leaf diseases in actual agricultural environments presents pragmatic difficulties concerning hardware needs, scalability, and computational efficiency. Real - time processing skills are still a major obstacle to end users' general adoption and implementation of these systems, especially in situations with limited computational capacity and internet connectivity. [18]

To improve the effectiveness, dependability, and accessibility of automated disease detection systems in agriculture, it is imperative to address the drawbacks and restrictions in current methods of leaf disease categorization utilizing computer vision techniques. Important issues for future study and innovation include overcoming leaf appearance fluctuation, strengthening model interpretability, minimizing overfitting, improving dataset quality and variety, and maximizing real - time processing capabilities. By tackling these issues, scientists and industry professionals can create leaf disease classification methods that are more reliable, scalable, and easy to use, thereby enhancing crop health, productivity, and global food security.

## 4. Method

The classification of leaf diseases has undergone a revolution thanks to machine learning and deep learning algorithms, which provide strong tools for automated picture analysis and pattern identification. The development of precise and scalable methods for identifying, treating, and managing leaf diseases in agricultural crops is made possible by these algorithms. This paper presents an overview of the many machine learning and deep learning algorithms used in the categorization of leaf diseases, emphasizing their benefits, applications, and methods in agriculture.

### 4.1 Machine Learning Algorithms

- **Support Vector Machines (SVM):**

SVM is a supervised learning algorithm that uses the optimal hyperplane to divide data points into distinct classes in order to classify the data. SVMs have been widely employed in the classification of leaf diseases to discriminate between healthy and sick leaves using handmade features (such as texture, color, and shape descriptors) derived from leaf photos [11]. SVMs can be challenging to use with huge datasets and nonlinear separable classes, but they do have good classification accuracy and noise resilience.

- **Decision Trees:**

Recursively dividing the feature space according to the most discriminative traits is what decision trees, which resemble trees, do. Decision trees can handle both numerical and categorical data and are capable of capturing complex decision boundaries in the classification of leaf diseases. Nevertheless, decision trees may not be able to generalize to previously untested data because to their propensity for overfitting, particularly with deep and complicated trees.

- **Random Forests:**

Several decision trees are combined in random forests, an ensemble learning technique, to increase classification robustness and accuracy. Random forests have been demonstrated to perform better in the categorization of leaf diseases than individual decision trees because they minimize overfitting and capture a wide range of data. Large - scale leaf disease datasets can benefit from random forests since they are less susceptible to noise and outliers and are computationally efficient.

### 4.2 Deep Learning Algorithms

- **Convolutional Neural Networks (CNNs):**

CNNs are deep learning architectures that draw inspiration from how animals' visual cortexes are organized. Convolutional, pooling, and fully connected layers make up the numerous layers of CNNs, which automatically extract hierarchical features from unprocessed pixel input [12]. CNNs have outperformed conventional machine learning algorithms in the categorization of leaf diseases by efficiently identifying spatial interdependence and hierarchical structures in leaf pictures. Transfer learning strategies improve classification accuracy and generalization capacities even more. One such strategy is fine - tuning pre - trained CNN models on datasets related to leaf diseases.

- **Recurrent Neural Networks (RNNs):**

Deep learning architectures called RNNs are made to handle sequential data that has temporal relationships [17] [19]. While RNNs are less frequently utilized in the classification of leaf diseases than CNNs, they can be applied to time - series data or sequential picture data, including films showing the growth and spread of diseases on leaves. RNNs are useful for evaluating the spatiotemporal dynamics of leaf diseases because they can capture long - term relationships and dynamic patterns in temporal sequences [13] [14].

Precision farming, yield prediction, crop disease detection, and other agricultural activities have all been addressed by machine learning and deep learning algorithms. These algorithms facilitate automated disease diagnosis, monitoring, and management in leaf disease classification, giving farmers timely insights and useful information for crop protection. Researchers and practitioners can create scalable and affordable solutions to solve the problems of leaf diseases and improve agricultural productivity and sustainability by utilizing cutting - edge computational approaches [15].

### 4.3 Proposed model

For the model, the researcher took pictures of a maize crop using a DJI Tello drone. The first step in the procedure is to take pictures of leaves. Python makes image processing tasks like image reading, preprocessing, and feature extraction easier using packages like OpenCV. The content of these photos is then analyzed and understood using computer vision techniques, which include finding patterns or characteristics suggestive of diseases.

To employ color segmentation and defragmentation to assess the health of the plants. The various color detections—red, brown, yellow, and green—were implemented using the OpenCv color detection module.

The taken picture was put into memory and processed so that the system could distinguish between healthy and diseased plants.

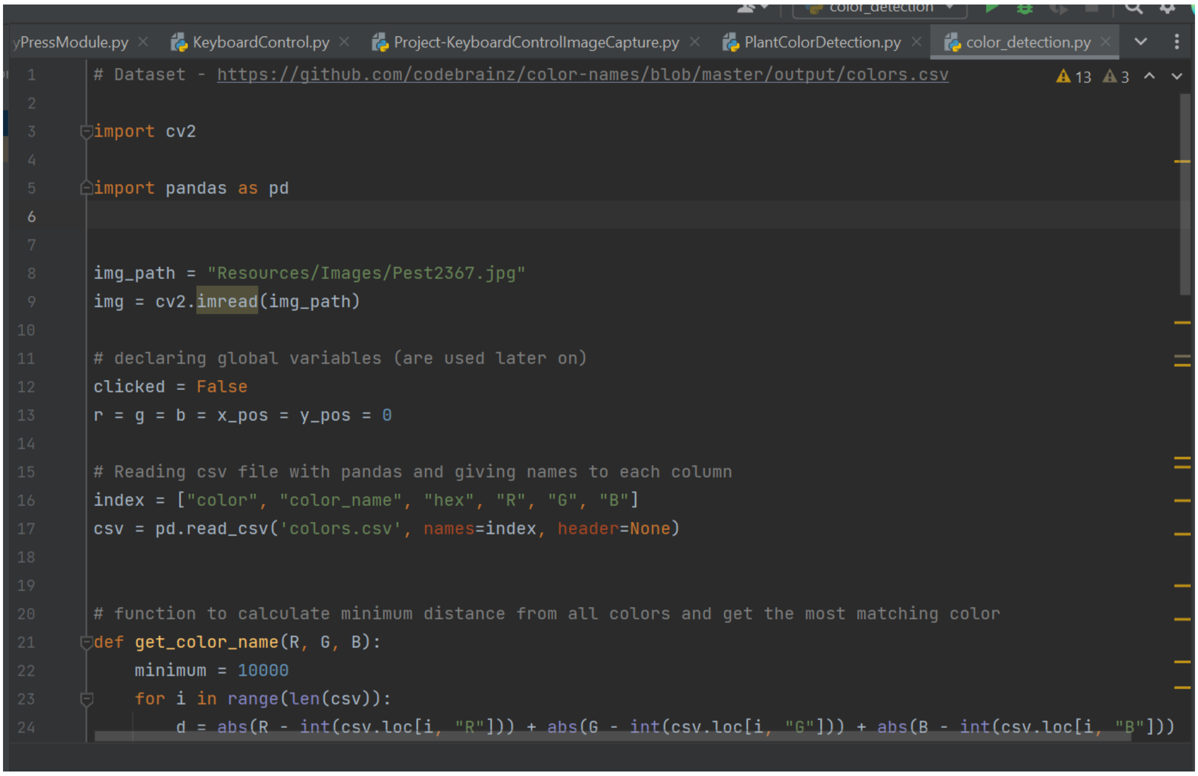


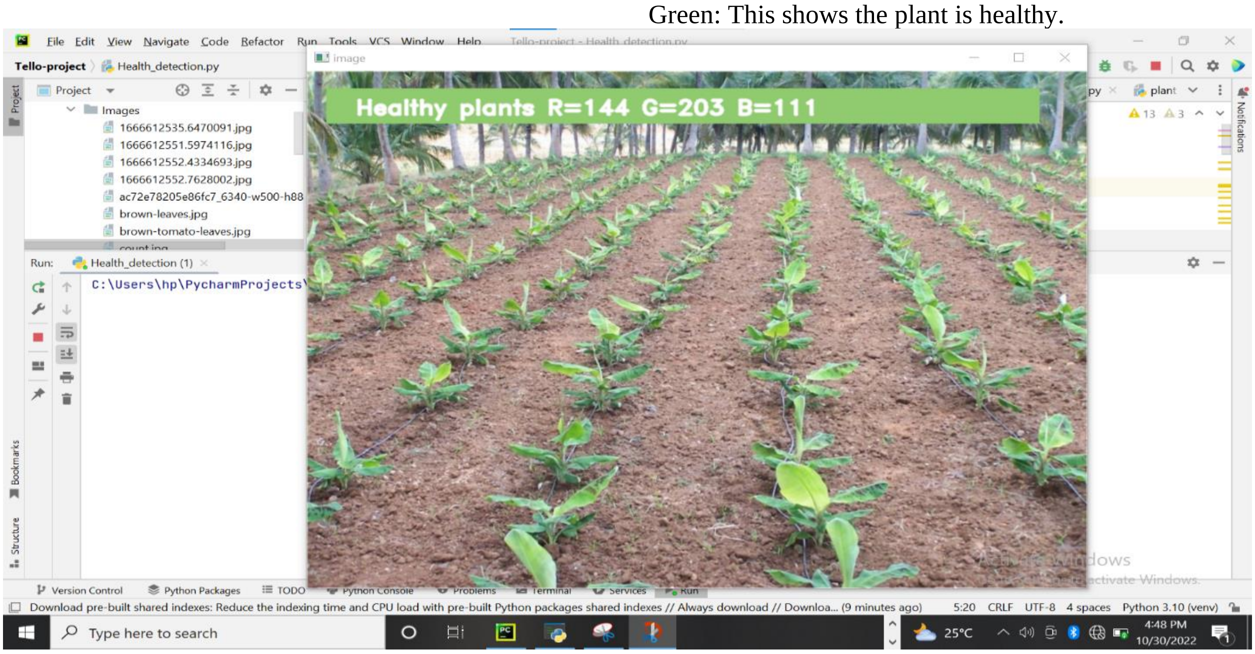

**Figure 1:** Python Model

The model loads the image first, then transforms it into a color space that highlights plant characteristics, like HSV (Hue, Saturation, Value), which is frequently more useful for analyzing color in plant leaves. In addition, the model establishes thresholds to assist in identifying leaf segments that exhibit peculiar hues that could potentially signify illness. These thresholds are used to build masks that are then used to isolate and examine the affected areas. In order to visually check the areas that were detected, the model would present the original image next to the mask. This appears in code as seen in the above figure:

## 5. Result

To use color segmentation and defragmentation to assess the health of the plants. The various color detections—red, brown, yellow, and green—were implemented using the OpenCv color detection module. The image captured was loaded in memory to be processed in other for the system to detect healthy plants from unhealthy plants.

The image above shows results of the health status of the plantain crops on a plantain plantation through image processing with the system.
Red or brown: This means the plant is dead
Yellow: This means the plant is unhealthy
Green: This shows the plant is healthy.





Paper ID: SR24601192902 DOI: https://dx.doi.org/10.21275/SR24601192902

## 6. Conclusion

Tools for classifying leaf diseases are provided by machine learning and deep learning algorithms, which enable automated plant disease detection and diagnosis in agricultural crops. Many different algorithms, each with pros and cons, have been used for leaf disease classification tasks, ranging from deep learning architectures like CNNs and RNNs to more conventional machine learning techniques like SVMs and decision trees.

The goal of this work was to develop a model for disease identification in corn fields by integrating Python Open Cy with color codes for disease diagnosis. Additionally, the researcher collected photographs for the core data set using a DJI Tello drone. The model performed well when it came to classifying maize illnesses.

As technology develops and agricultural concerns shift, there is great potential for improving crop health, productivity, and food security worldwide through the further application of machine learning and deep learning algorithms to agricultural practices.

## References

[1] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image - based plant disease detection. Frontiers in Plant Science, 7, 1419.
[2] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. Computational Intelligence and Neuroscience, 2016.
[3] Barbedo, J. G. A. (2019). Factors influencing the use of deep learning for plant disease recognition. Biosystems Engineering, 180, 80 - 91.
[4] Toda, Y., Honda, K., Ohsawa, R., & Takemura, Y. (2016). Automatic detection of diseases in persimmon leaves by deep learning and the feasibility of its implementation on a smartphone. Computers and Electronics in Agriculture, 130, 96 - 107.
[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp.1097 - 1105).
[6] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks?
[7] Cortes, C., & Vapnik, V. (1995). Support - vector networks.
[8] Breiman, L. (2001). Random forests.
[9] T., & Hart, P. (1967). Nearest neighbor pattern classification.
[10] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.
[11] Hochreiter, S., & Schmidhuber, J. (1997). Long short - term memory.
[12] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine.
[13] Sabour, S., Frosst, N., & Hinton, G. E. (2017). Dynamic routing between capsules.
[14] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.
[15] Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks.
[16] Sabour, S., & Hinton, G. E. (2017). Capsule networks for low - data transfer learning.
[17] Chiadika OD, Li M, Choi J. Improving training time in capsule neural network. In2021 3rd International Congress on Human - Computer Interaction, Optimization and Robotic Applications (HORA) 2021 Jun 11 (pp.1 - 4). IEEE.
[18] Mantey EA, Zhou C, Anajemba JH, Okpalaoguchi IM, Chiadika ODM. Blockchain - secured recommender system for special need patients using deep learning. Frontiers in Public Health.2021 Sep 20; 9: 737269.
[19] Onyeachonam Dominic - Mario Chiadika, Oghenerukevwe Affun, Ujerekre Ekoko, Ochuko Goodluck Utu, 2024, A Comparison of Artificial Intelligence Methods for Predicting the Rate of Violent Crime, International journal of engineering research & technology (ijert) volume 13, Issue 04 (April 2024),