International Journal of Computing and Artificial Intelligence

E-ISSN: 2707-658X
P-ISSN: 2707-6571
www.computersciencejournals.com/ijcai
IJCAI 2025; 6(1): 123-132
Received: 24-01-2025
Accepted: 03-03-2025

Mohammad Anwar Hossain
Department of Computer Science & Engineering, World University of Bangladesh, Dhaka, Bangladesh

Newaz Ibrahim Khan
Department of Computer Science & Engineering, World University of Bangladesh, Dhaka, Bangladesh

Noshin Un Noor
Department of Computer Science & Engineering, World University of Bangladesh, Dhaka, Bangladesh

Fazle Rabby
Department of Computer Science & Engineering, World University of Bangladesh, Dhaka, Bangladesh

Ashiqur Rahman
Department of Computer Science & Engineering, World University of Bangladesh, Dhaka, Bangladesh

Hridoy Chandra Shill
Department of Computer Science & Engineering, World University of Bangladesh, Dhaka, Bangladesh

Corresponding Author:
Mohammad Anwar Hossain
Department of Computer Science & Engineering, World University of Bangladesh, Dhaka, Bangladesh

Crop health analysis system: Integrating machine learning for disease detection in agricultural images

Mohammad Anwar Hossain, Newaz Ibrahim Khan, Noshin Un Noor, Fazle Rabby, Ashiqur Rahman and Hridoy Chandra Shill

DOI: https://www.doi.org/10.33545/27076571.2025.v6.i1b.141

Abstract
The proposed project is a pioneering research effort utilizing cutting edge technology to revolutionize plant disease detection within agriculture. Utilizing CNN for feature extraction and ResNet-50 for classification, this system outperforms in the accuracy of providing crop disease from the agricultural images. This approach utilizes a formal Image Classification Cycle, which offers processes for data acquisition, augmentation or data processing, thereby improving the accuracy and efficiency of detection. Novel features entail merging sophisticated machine learning models with a simple and clean web application via the Streamlit framework allowing on-line disease identifying for farmers. When comparing the accuracy and performance of the system against the leading ResNet-50 algorithm, it is noted that the accuracy is 95% with optimized ResNet-50 achieving 98.72% accuracy. This implementation combines the speed, accuracy, and scalability of manual observation and traditional algorithms. It also discusses functional requirements, such as dataset reliability, and non-functional properties, like usability, robustness, and adaptability to a diverse agricultural condition. Finally, this work is distinguished from other solutions by its speed, scalability for small and large farms, and flexibility to environmental challenges. The system, which can be used by farmers as a tool, is expected to increase crop yield, minimize dependence on pesticides, and promote sustainable agricultural practices. The results highlight the capacity of machine learning powered systems to change agriculture and support global food security.

Keywords: Crop disease detection, convolutional neural networks (CNNs), resnet-50, image classification cycle, real-time diagnosis, agricultural technology

1. Introduction

1.1 Research Background

The agricultural industry is addressing some of the most pressing challenges, such as how to keep our crops healthy and productive. Certainly, crop diseases caused by pathogens such as fungi, bacteria, viruses, and nematodes are a serious threat to global food security as they lower crop yield and lead to severe economic losses to farmers. In addition, these diseases often require more extensive use of chemical treatments, which can damage the environment. Addressing these issues is critical to promoting sustainable agricultural practices and protecting food supply chains.

Crop disease detection is done through manual observation which consumes a lot of time, is error-prone and subjective. On the other hand, recent advances in machine learning (ML) provide a game-changing approach. With ML algorithms trained on vast datasets of agricultural images, these systems can detect disease patterns like no human ever could, and with far greater speed. Our method combines CNNs for features extraction and ResNet-50 for classification in order to improve the detection of diseases with regard to both accuracy and fastness compared to traditional methods.

Moreover, CNNs are quite effective at identifying complex spatial structures in images, which is particularly useful in determining nuanced disease characters in plants. One of the previously mentioned cutting-edge deep learning architectures, ResNet-50, relies on this principle while leveraging its powerful residual learning blocks to provide very accurate predictions. All these models work synergistically to develop a system that adapts to the complexities of evolving crop diseases.

The main objective of this project is to build an advanced crop health analysis model using

ML algorithms for precise and early disease detection. Our objective is to give farmers real-time data regarding crop development so they can take productive steps towards propelling themselves down this path specifically, by developing this system as an easy-to-use web application.

1.2 Objectives

- To design a crop health analysis system using machine learning algorithms capable of accurately detecting and classifying plant leaf diseases from images.
- To develop a web application that seamlessly integrates the trained ML model, allowing users to upload images and receive instant disease diagnoses.

1.3 Scope of study

In this paper, we present a novel system that utilizes machine learning for the detection of crop disease through image processing. Key components include:

- **ML Models:** Creation of custom-built ML Models to identify illness of crops via image processing methods.
- **Performance evaluation:** Testing performance on real-world datasets using measures like accuracy, precision, and recall.
- **System Integration:** Integrating with other farm analytics systems, these systems need to be integrated with existing farming systems to give practical recommendations that can help the farmers make effective decisions.

2. Related Works

The rapid advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have revolutionized image classification tasks across various domains, including agriculture. CNNs excel at extracting intricate spatial features from images, enabling precise predictions. Models like VGG-16, InceptionV3, and EfficientNetB7 have demonstrated exceptional accuracy on benchmark datasets such as ImageNet. For instance, EfficientNetB7 outperformed VGG-16 and InceptionV3 in a study on butterfly and spider classification, achieving an impressive 99% accuracy compared to 97.67% and 97.2%, respectively [1].

In the context of agricultural disease detection, significant progress has been made in leveraging machine learning techniques. A study focused on rice leaf disease detection in Bangladesh, one of the top ten rice-producing countries, highlights the efficacy of automated systems over manual methods due to time and labour constraints. The study targeted three common rice diseases: leaf smut, bacterial leaf blight, and brown spot. By utilizing clear images of affected rice leaves and applying pre-processing techniques, the study trained various machine learning algorithms, including K-Nearest Neighbour (KNN), Decision Tree (J48), Naive Bayes, and Logistic Regression. Notably, the Decision Tree algorithm, after 10-fold cross-validation, achieved an impressive accuracy of over 97% on the test dataset, underscoring the potential of machine learning in enhancing agricultural disease management [2].

Conventional approaches are labor-intensive and dependent on domain expertise. Recent advancements in deep learning-to identify diseases from digital images-offer potential solutions. But the large parameter size of deep learning models poses challenges. An end-to-end deep learning model with built-in hyper-parameter optimization was proposed for healthy and unhealthy corn plant leaves in a recent study. Using pre-trained CNNs such as EfficientNetB0 and DenseNet121, and performing data augmentation, the classification accuracy reached an extraordinary 98.56%. The effectiveness of the proposed method was further validated by comparative analysis with other pre-trained CNN models. The study underscores the potential of deep learning techniques to revolutionize the automation of plant disease diagnostics, while also noting the critical need for effective parameter optimization to further improve efficiency and scalability [3].

Chili, where its delicious and spicy fruit is eaten worldwide, is one of the most cultivated crops that needs a precise health diagnosis for maximum yield. Deep learning has demonstrated potential in recognizing plant disease; however, it requires a large amount of data, which is expensive to obtain. In order to solve this, A particularly small dataset of healthy and diseased chili leaves, had been augmented with the use of geometric transformations. While CNN and ResNet-18 models showed 97% average accuracy on augmented datasets [4].

In agricultural economies such as India, where a majority of the population relies on agriculture, early detection of plant leaf diseases is critical. Early identification of leaf diseases is very important to avoid loss and to ensure food security. In this study, an approach based on image processing, including image segmentation, clustering and some open-source algorithms was designed for disease identification in tomato plant leaves. The purpose of these techniques is to develop a dependable, safe and accurate method of tomato plant leaf disease detection [6].

A. Shaikh *et al*. explores the impact of digitalization and ICT in agriculture, highlighting benefits for farmers and consumers. It covers the roles of robotics, IoT devices, machine learning, AI, and sensors in modern farming. Additionally, it examines the use of drones for crop observation and yield optimization. The paper reviews recent literature and discusses current and future trends in AI in agriculture, identifying emerging research challenges [5, 7].

Pushpa *et al*. presents a model for detecting and identifying crop leaf diseases using the CNN- based AlexNet model. Compared to VGG-16 and Lenet-5, AlexNet showed superior accuracy. Using a dataset of 7070 images from the Plant Village repository, the model achieved a 96.76% accuracy rate in identifying various diseases in corn, rice, and tomato plants [8].

Previous studies developed a method to identify leaf diseases in tomato plants, enhancing classification accuracy and reducing computational time. They utilized color histograms, Hu Moments, Haralick, and Local Binary Pattern features for training and testing. The study compared random forest and decision tree classifiers, finding that the random forest achieved a higher accuracy of 94% compared to 90% for the decision tree [9].

Tomatoes rank among the foremost horticulture crops globally, yet Early Blight poses a substantial threat to their yield in India. Recognizing the urgency in timely identification, we explore the classification of Early Blight in tomato leaves using ResNet and Xception networks, achieving a remarkable accuracy of 99.952% with self-collected images. Despite existing efforts in disease classification, the lack of spatial information for affected leaves steers our focus towards object detection. Leveraging

YOLO framework variants, we prioritize both accuracy and real-time inference, aiming to support the development of a mobile application for disease identification [10].

Amidst soaring population growth, achieving a twofold increase in global crop productivity by 2050 is imperative. However, pests and diseases present formidable barriers to this goal. To address this challenge, efficient methods for automatic detection, identification, and prediction of agricultural pests and diseases are crucial. Machine Learning (ML) techniques offer promising solutions by extracting insights from agricultural data. The authors conduct a literature review on ML techniques in agriculture, with a focus on disease and pest management, particularly in tomato crops. By advancing smart farming and precision agriculture, the survey aims to empower farmers to minimize pesticide use while enhancing crop quality and productivity [11].

Detecting chili diseases early is crucial due to their detrimental impact on plant health and yield. This paper explores an efficient method for early disease detection using leaf features inspection. By capturing and processing leaf images, the health status of each plant can be determined. This approach ensures targeted application of chemicals, optimizing resource usage. Leveraging image processing techniques, hundreds of chili disease images can be analysed, providing a cost-effective solution to assist farmers in monitoring large plantation areas [12].

Accurate plant disease detection is vital for agricultural sustainability. The researchers focus on tomato, chili, potato, and cucumber diseases, proposing an AI-driven approach for prediction. It reviews ML and DL models, outlines disease symptoms, and discusses challenges and future prospects. By integrating AI with IoT platforms like smart drones, it aims to enhance field-based disease monitoring. This analysis assists practitioners in navigating plant disease detection strategies and understanding current limitations [13].

3. Methodology

A methodology is a systematic approach to solving a problem or achieving a goal. It is a set of principles and procedures that are used to guide the research process. For the proposed research, the Image Classification Cycle has been selected as shown in figure 1.

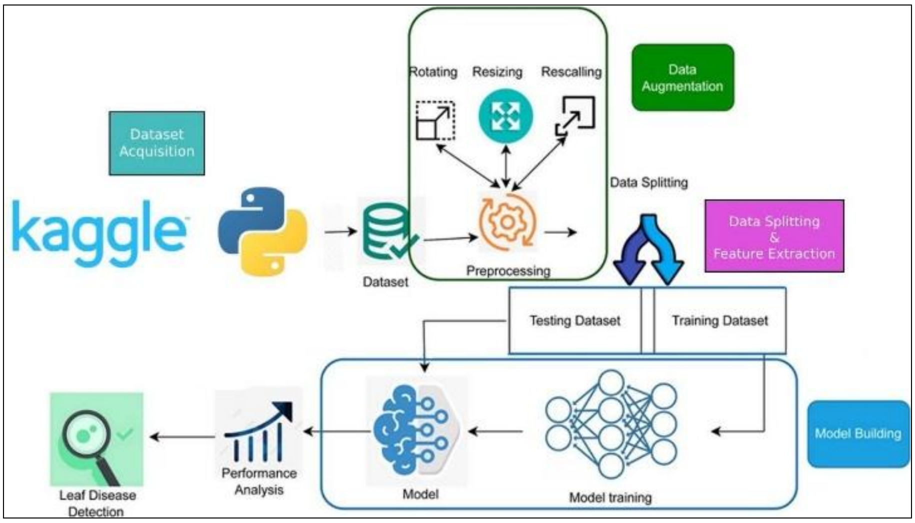

Fig 1: Image Classification Cycle (Overview of proposed methodology)

This Image Classification Cycle, a framework that provides us a structured approach to classify several images for multiclass detection. This methodology encompasses several key stages and each stage is designed to ensure a comprehensive and organized response to crop disease detection

3.1 Description of Methodology

The proposed project work, titled "Crop Health Analysis System: Integrating Machine Learning for Disease Detection in Agricultural Images," presents a structured methodology comprising six key phases to facilitate accurate and efficient leaf disease detection in agricultural images. The following covers the details about the methodology chosen:

3.1.1 Dataset Acquisition: For our Crop Health Analysis System, we procured a dataset from Kaggle, a platform known for its diverse collection of datasets and machine learning resources. The dataset encompasses a wide range of agricultural images depicting various crop diseases along with healthy plant samples. This dataset forms the foundation of our research, providing the necessary examples for training and evaluating our machine learning models.

3.1.2 Dataset Augmentation: We used data augmentation, which we used to improve the variability and quality of the dataset. Their pre-processing involved rotating, resizing, and rescaling of images. This image augmentation strategy was adopted to enrich the data used in the training phase and to increase the robustness of our models to different environmental conditions that can occur in most agricultural crops in real life.

3.1.3 Data Splitting and Feature Extraction: After the pre-processing we split the dataset into a training and testing subsets to train our models on an array of diverse examples while keeping it out of the testing data. Much of that was spent on training, allowing the models to learn good representations of crop diseases. The testing subset is used to evaluate the generalisation performance of our models. CNNs were used after that to perform feature extraction to identify critical patterns from the training data.

3.1.4 Model Building: During model building phase, we utilized dual approach using Convolutional Neural Networks (CNN) for feature extraction, and ResNet-50 for detection of image classes. We apply this approach to benefit from the representational ability of CNNs appropriate for the notable discriminatory features of agricultural images, followed by ResNet-50 for classification to improve detection amongst different diseases.

3.1.5 Performance analysis: We performed comprehensive performance analysis following model training to analyze the efficacy of our models. The metrics such as accuracy, precision, recall and F1-score were calculated for the models to evaluate crops diseases detection performance. Also, we used visual inspection of the model predictions to have an understanding of their behavior towards real-world agricultural images. Such analysis served as useful insight for gradationally refining the models and optimizing their performances.
By rigorously adhering to this methodology, our project work aims to contribute to the development of robust and efficient solutions for crop health analysis, ultimately enhancing disease detection and management in agriculture.

3.2 Requirement analysis
In this section, we analyses the essential requirements and objectives of our study, which centers on developing a Crop Health Analysis System integrating machine learning for disease detection in agricultural images. Our approach focuses on detecting crop diseases and providing actionable insights to enhance agricultural productivity and crop health management.

3.2.1 Functional Requirements
Functional requirements are those that define the essential actions and functions a system must carry out to achieve its designed purpose. Functional requirements specify how the system should function in terms of using machine learning techniques to detect and analyse crop diseases and this is important in this research.
The functional requirements of the proposed work are as below:

- **Dataset:** A sound data source is a major requirement for this research. We have used a dataset available in kaggle that contains lots of images of healthy and diseased crops. This data set is important for our research as it gives a ground evidence based analysis of crop disease.
- **Python Programming Language:** Python serves as the foundation for our research, enabling all stages of the development of our crop health analysis system. This powerful and flexible programming language can be used for pre-processing, training a model, testing it, and more.
- **Computing and Technical Environment:** Our research explores the far-reaching aspects of having a solid and secure computing environment to carry out the experiments successfully. A hihg performance computer system is needed for such resource-intensive ML tasks that is compatible with a collaborative research environment like Google Colab.
- **Python Libraries:** In our research we make extensive use of some powerful Python libraries that serve our purpose of fulfilling the system requirements. To further streamline our project, we relied on an array of libraries, from Pandas for data manipulation and NumPy for numerical operations to TensorFlow for implementing our deep learning model. Python has a rich ecosystem of libraries that provide us powerful tools in data pre-processing, feature extraction, model training, and evaluation. We bring this requirement up in order to highlight how fitting these libraries into our research ensures that the system is appropriate and performs properly within a practical agricultural disease detection environment.
- **Machine Learning Algorithms and Classifiers:** We leverage different machine learning algorithms and classifiers (in our machine learning research), which help us reinforce the crop health analysis system. We start with CNN (Convolutional Neural Networks) which are expert in discovering significant patterns in our data thus a good option for feature extraction. This process utilizes ResNet-50, a deep residual network that excels in intricate image classification tasks. Our integrated system will thus detect crop diseases effectively using these advanced models.
- **Web Application Interface:** The system has then to be made accessible and user-friendly, hence leading us to develop a web application interface using which the farmers and the agricultural experts can upload images and get instant analysis. The interface is intended to be intuitive and straightforward.
- **Performance:** Performance is the system's capacity to operate its tasks effectively and provide timely responses.
- **Accuracy:** Accuracy is the ability to generate correct and reliable results/outputs.
- **Security:** Security policies and strategies protect data, systems and users from threats and intrusions.

Hence, the functional requirements of our research form the backbone of our endeavor to enhance crop health and tackle the challenge of disease detection through machine learning techniques.

3.2.2 Non-Functional Requirements
Non-functional requirements focus on qualities beyond what the system is supposed to achieve, which is crucial in an agricultural real-world context. It is important in our project that our crop health analysis system for disease detection is effective, reliable, and practical.
The non-functional requirements of the proposed work are:

- **Robustness:** Robustness means the system's tenacity for its core functions through adverse conditions or attacks.
- **Usability:** It is defined as the ease with which the users are able to interact with the system and how they can use it.
- **Interpretability:** The system needs transparency so that results are comprehensible for users, which contributes to the making informed decisions and taking timely actions.
- **Flexibility:** The system needs to be flexible to different agricultural conditions, crop variety and geographies so as to be meaningful and applicable in different settings.

In addition to the functional requirements, the successful implementation of machine learning techniques to tackle crop disease detection should also priorities non-functional requirements, ensuring the system can effectively adapt to evolving agricultural challenges and maintain high standards of crop health management.

3.3 Design of Proposed work

Here, we present the design of our system, structured into two key sections. The first, "Machine Learning Model for Crop Disease Detection," details the methodology and framework behind our model, emphasizing algorithmic foundations and specifics for agricultural image analysis. The second, "Workflow of Crop Health Analysis System," outlines how data pre-processing, model training, and validation ensure effective disease detection and crop health management.

3.3.1 Machine Learning Model for Crop Disease Detection

In the flowchart of Figure 2, the process starts with raw data, the initial step involves rigorous pre-processing to enhance its quality and prepare it for subsequent stages. The dataset is then partitioned into distinct training and testing subsets to facilitate effective model training and evaluation. Utilizing Convolutional Neural Network (CNN) techniques, feature extraction is performed on the training data to capture pertinent information crucial for the task at hand.

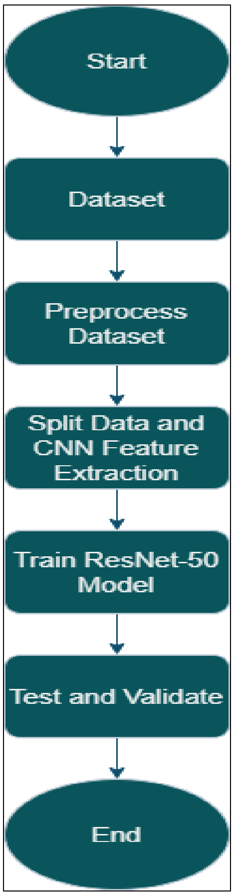

Fig 2: Proposed Machine Learning Model

Subsequently, a ResNet-50 model is trained using these extracted features, leveraging its deep learning capabilities to learn and generalize patterns within the dataset. Finally, the model's performance is rigorously assessed and validated using the independent testing set, ensuring robustness and reliability in its predictive capabilities for real-world applications. This systematic approach ensures that each phase from data pre-processing to model evaluation is meticulously executed to achieve optimal results in data-driven analysis using advanced neural network architectures.

3.3.2 Workflow of Crop Health Analysis System

Here, we present the design of our Crop Health Analysis System, which is encapsulated in a detailed flowchart to illustrate our approach. The flowchart, depicted in Figure 3, serves as a fundamental visual representation of the system's workflow, highlighting the crucial steps involved in disease detection and user notification.

Using a diagram, it represents the entire process, to get an idea of how the system works, (From uploading an image to getting a disease diagnosis). As illustrated in Figure 3, the first step is to upload an image of crop leaf via web application interface. The trained model classifies the uploaded image, using a CNN for feature extraction and ResNet-50 for image classification. We trained this model to predict various crop diseases from the uploaded images.

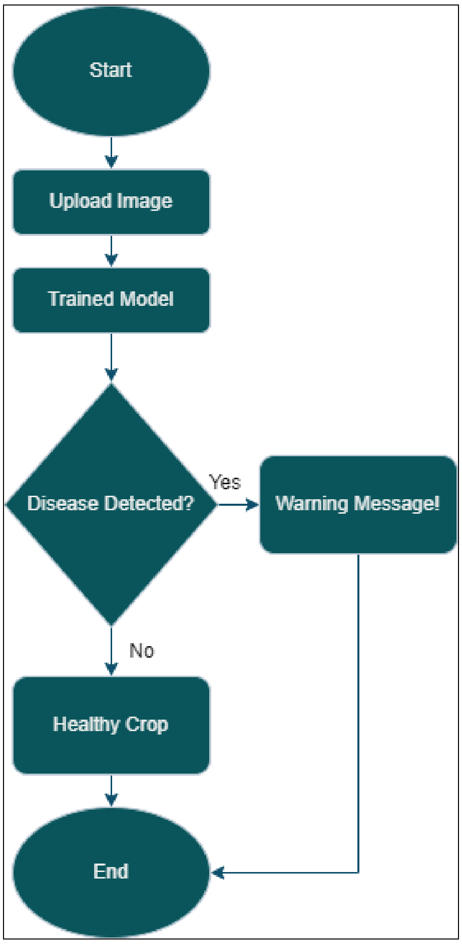

Fig 3: Workflow of Crop Health Analysis System

After the image is analyzed, the system will identify if a disease is detected:

If any disease is detected, then the system gives a warning message which shows disease name that was detected and the user is warned to take further necessary steps.

In the absence of disease detection, it shows healthy crop message which shows that the crop is healthy.

This streamlined process ensures rapid and accurate disease detection, offering farmers vital insight into managing their crops efficiently.

3.4 Development of Proposed work

We will provide a comprehensive account of the development process, transitioning from the theoretical framework to the practical realization of our objectives here:

3.4.1 Pre-Processing

This section describes the process followed in preparing our dataset, which is obtained from Kaggle. Data augmentation was performed on the datasets so that different operations like rotation, resizing, and rescaled were performed on the datasets to increase the variability of the dataset. This phase made sure that the data is clean, structured, and ready for further analysis.

3.4.2 Cross-Validation and Feature Selection

The dataset was divided into training and test subsets after pre-processing to bring up a model for training and testing purpose. Next step was using Convolutional Neural Networks (CNNs) to extract features from the training dataset. Using well-known CNN architecture, the important patterns and features from the images were extracted that was important for disease classification.

3.4.3 Training & Testing ResNet-50 Model

After the initial feature extraction process, we trained/tested the model over ResNet-50 architecture. We trained our model on the extracted features using ResNet-50, a convolutional neural network architecture that is robust for an image classification task. This model's assessment was done in terms of different precision, recall, F1 score, which is very crucial for determining whether a model can detect crop diseases correctly.

3.4.4 Web App Integration

The final stage for us was to build a web application to serve the trained model. This includes a web app that has a clean and straight-forward UI for users to upload crop images. The uploaded image is then processed through the trained model by the integrated system, which detects diseases and displays the results in real-time. In case any disease is detected it returns a warning message with name of the disease otherwise returns healthy.

4. Project Description and Result Analysis

4.1 Project Description

In this section, we will provide few sections, some gives the screenshots with descriptions of development processes, which fulfil the objectives and justification of our research. And also shows results, which proves the reliability and adaptability of our system.

4.1.1 Downloading and Unzipping the Dataset

Downloading a dataset is a fundamental and essential step in the process of data analysis, machine learning and data-driven research.

```
[ ] !pip install -q kaggle

[ ] !mkdir ~/.kaggle

[ ] !cp /content/drive/MyDrive/CropDisease/kaggle.json ~/.kaggle/

[ ] !chmod 600 ~/.kaggle/kaggle.json

[ ] !kaggle datasets download vipoooool/new-plant-diseases-dataset

    Dataset URL: https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset
    License(s): copyright-authors
    Downloading new-plant-diseases-dataset.zip to /content
    100% 2.70G/2.70G [00:17<00:00, 146MB/s]
    100% 2.70G/2.70G [00:17<00:00, 164MB/s]
```

Fig 4: Dataset Downloading Code and Output

Figure 4 shows the dataset downloading, which typically involves using libraries, such as kaggle in Python, to download dataset from kaggle.

Unzipping a dataset is a critical initial step to extract its contents from a compressed archive. This process allows access to the dataset's files, enabling data exploration, analysis, and machine learning model training.

```
!unzip new-plant-diseases-dataset

  inflating: new plant diseases dataset(augmented)/New Plant
  inflating: new plant diseases dataset(augmented)/New Plant
  inflating: new plant diseases dataset(augmented)/New Plant
  inflating: new plant diseases dataset(augmented)/New Plant
  inflating: new plant diseases dataset(augmented)/New Plant
  inflating: new plant diseases dataset(augmented)/New Plant
  inflating: new plant diseases dataset(augmented)/New Plant
  inflating: new plant diseases dataset(augmented)/New Plant
```

Fig 5: Dataset Unzipping Code and Output

Figure 5 demonstrates the unzipping process, essential for preparing the dataset for further use. This step is typically performed to decompress files downloaded from online repositories or archives.

4.1.2 Pre-process & CNN Feature Extraction

For this project, we used CNNs to extract features from images in a dataset of plant diseases. First, we started with preprocessing the dataset, which included resizing the images and preparing data generators using Image Data Generator from TensorFlow. We then went ahead to build a CNN model which consists of convolutional/pooling layers and then fully connected layers for classification.

Once we have trained and saved this model, we went ahead to extract features from the trained CNN via the 'flatten' layer that converts the learned representation for each image into a feature vector. The extracted features, which are the individual properties of the images learned by the convolutional neural network, were stored for subsequent model training and analysis. Such feature extraction allows for the previously learnt representations to be reused in transfer learning via other models such as ResNet-50, leading to better accuracy and reduced computational cost for the following training stages.

4.1.3 ResNet-50 Model Training and Validation

We used the extracted features to train and validate a ResNet- 50 model after the feature extraction step and using a CNN model trained on a datasets of photos of plant diseases. To do so, the extracted features, which were stored in numpy arrays after pre-processing and being able to model were reshaped to match the inputs of ResNet-50. A pretrained ResNet-50 model (on ImageNet) was loaded without its top classification layers. New layers were added on top of the ResNet-50 backbone for fine-tuning the model to classify plant diseases. The model was compiled using the Adam optimizer and categorical crossentropy as the loss function. Extracted features were used to train and validate the model, making use of already learned representations of images for efficient usage. This method utilizes transfer learning, which may enable a performance boost in classification, whilst lowering the computational cost of training from zero.

4.1.4 Integration with WebApp for Model Deployment

In order to deploy our machine learning model and enable interaction with it, we used a type of code called Streamlit that allows us to build web applications using Python. Using Streamlit, we created an intuitive user interface that enables easy input of data, visualization of predictions, and exploration of model insights. This includes interactive elements like sliders, selection dropdowns, and text inputs allowing for input parameters to the model. The real-time update feature from Streamlit allowed me to do the changes to models and receive live feedback for predictions. Additionally, it makes our machine learning solution easier to be adopted, as it does not require to go through the normal IT channel, keeping our machine learning solution usable for technical and non-technical users alike.

4.1.5 Web Application Interface for Crop Disease Diagnosis

The interface of the web application we built is the entry point of the advanced crop health analysis system deployed with the powerful machine learning model for disease diagnosis. Using Streamlit to deploy as an interface, users can upload their crop images directly. Our machine learning model processes these images in rapid succession, utilizing deep learning protocols to detect and classify crop diseases accurately from their visual properties as shown in figure 6.

Fig 6: Dataset Unzipping Code and Output

The interface not only enhances user accessibility but also ensures efficient and reliable disease identification. Through rigorous testing and validation, our system consistently achieves high accuracy in disease classification, validating the efficacy of our integrated approach in addressing agricultural challenges using state-of-the-art machine learning technologies.

4.2 Result Analysis

From the performance metrics displayed in Figure 7, it is evident that ResNet-50 outperformed both CNN and MobileNetV2 in our Crop Health Analysis System (CHAS). ResNet-50 achieved the highest accuracy of 0.95, precision of 0.95, recall of 0.95, and F1- score of 0.94, showcasing its robustness in classifying various crop diseases within our agricultural image dataset.

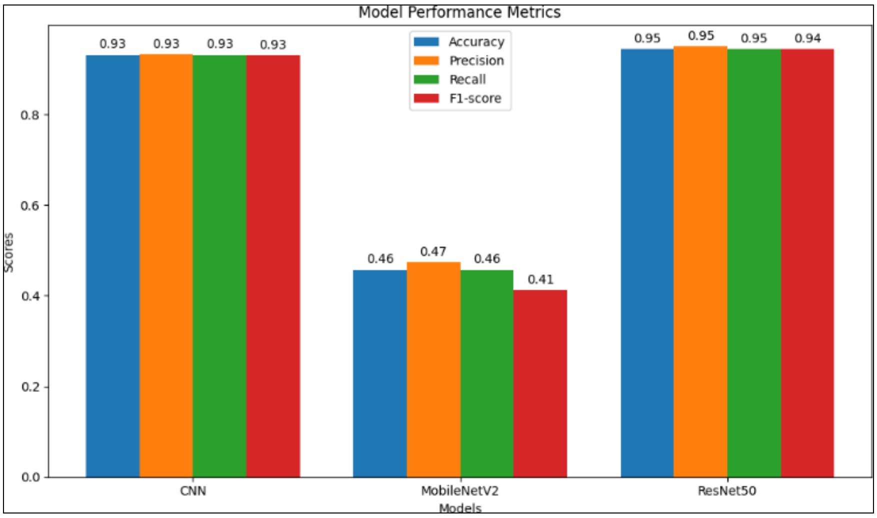

Fig 7: Comparison of Model Performance Metrics

In comparison, CNN exhibited an accuracy, precision, recall, and F1-score of 0.93, highlighting its competitive performance. On the other hand, MobileNetV2 demonstrated lower metrics with an accuracy of 0.46, precision of 0.47, recall of 0.46, and F1-score of 0.41. These results underscore ResNet-50's effectiveness as the preferred model for our system, providing farmers with a reliable tool for timely and accurate crop health management.

Moreover, our experimentation demonstrated the significant impact of incorporating CNN for feature extraction and training ResNet-50 using these extracted features in our Crop Health Analysis System (CHAS). This process notably enhanced ResNet-50's performance, achieving an accuracy of 98.72% in our final implementation. By leveraging CNN's capability to extract detailed features from agricultural images, we improved ResNet-50's ability to recognize subtle differences in crop health indicators. This optimization not only sped up the classification process but also provided farmers with a dependable tool for making accurate and timely decisions about crop health management.

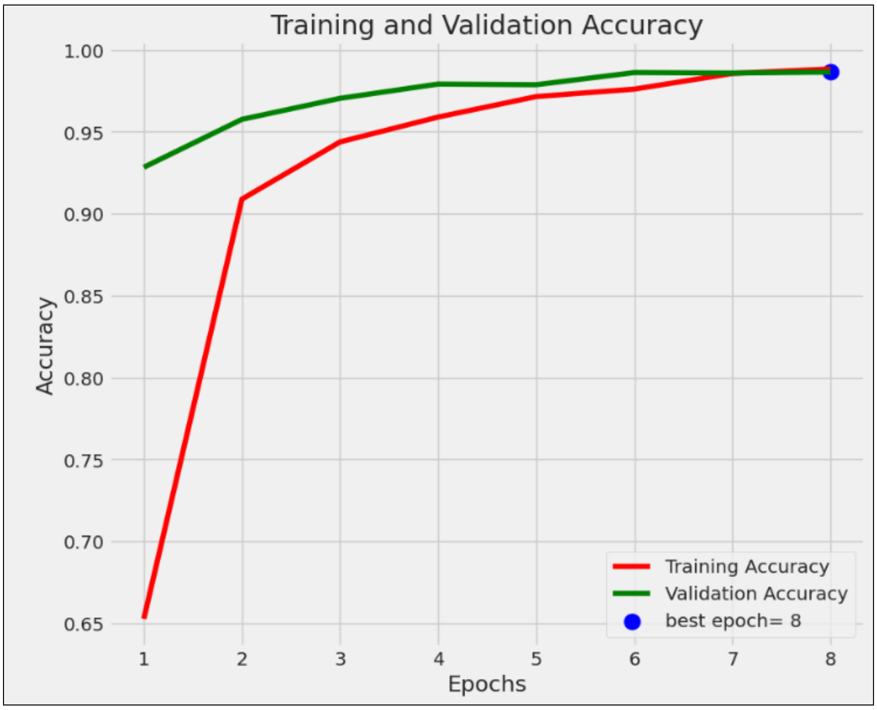

Fig 8: Training and Validation Accuracy

These findings are further supported by the training and validation metrics shown in Figures 8 and 9. Figure 8 depicts the training and validation accuracy, indicating how effectively our model learns and predicts outcomes on new data, crucial for evaluating its performance in real-world scenarios. Meanwhile, Figure 9 illustrates the training and validation loss, demonstrating our model's ability to generalize well to unseen data by minimizing loss during training.

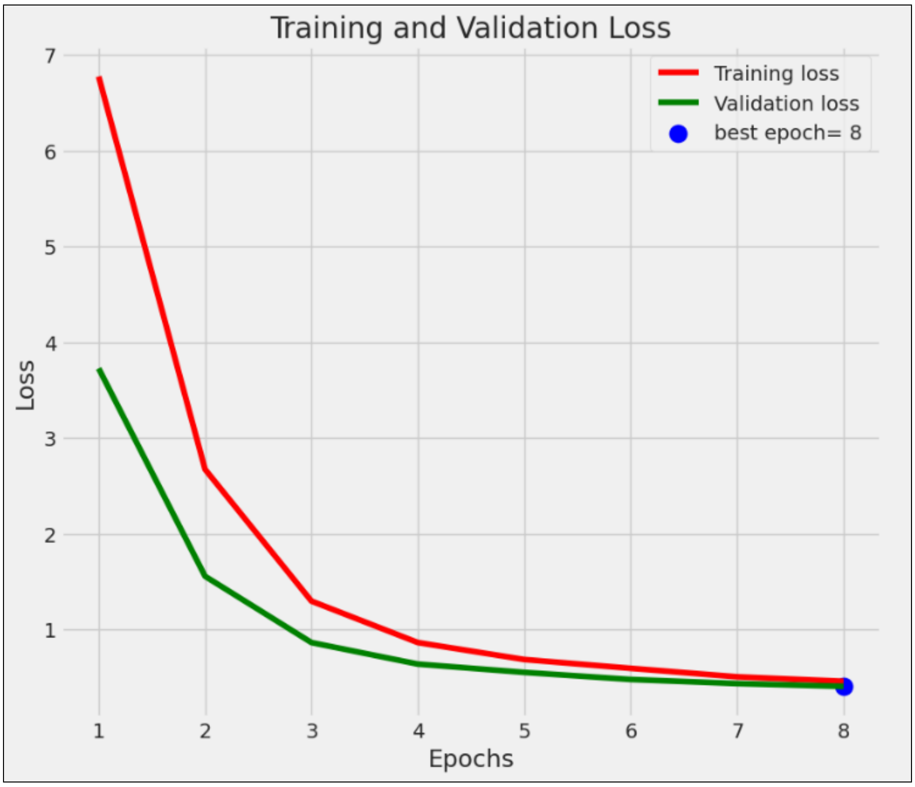

Fig 9: Training and Validation Loss

So the results shown here reflect the capabilities of our system for real-world applications. With the ability to provide an alternative to conventional disease detection methods and help farmers with crop health management decisions, our CHAS can support farmers by quickly diagnosing diseases more reliably and more efficiently. Additional research is required to assess the scalability of the system to larger farms and its compatibility with real-time mitigation strategies in production.

5. Conclusion

Our approach, the Crop Health Analysis System, leverages CNNs for basic feature extraction and ResNet-50 for final feature classification, and shows promising results compared to existing models in agricultural disease detection. With an optimized accuracy of 98.72%, this system outperforms traditional methods and existing state-of-the-art machine learning frameworks, such as MobileNetV2 and standard CNN architectures. The introduction of solid data augmentation methods and an organized Image Classification Cycle provides improved integration to various agriculture situations and environment-based cognition.

In contrast to the widely adopted labor-intensive traditional analytics that are often vulnerable to inaccuracies, this system employs deep learning for accurate, scalable, and real-time disease diagnostics. Furthermore, the realization of this work in terms of smooth integration of advanced machine learning algorithms into a convenient web application using the Streamlit framework differentiates it significantly, as it delivers real-time actionable insights to farmers. Such approaches not only enable farmers to make informed decisions but also encourage sustainable agricultural practices by reducing reliance on pesticides and increasing crop yield.

Performance comparison with all existing models establishes the robustness, cost-effectiveness, and effortless usage of this proposed system. Although prior research has achieved reputable outcomes, they frequently have constraints in scalability, computational cost, or suitability to real-life conditions. Through model architecture optimization and practical deployment strategies addressing these novel gaps, this research sets the next performance threshold for crop health analysis systems.

Lastly, the system we proposed is a revolutionary step in advancing agricultural technology with the best possible performance in detecting a disease. The new method can be directly adapted to different farming conditions while maintaining the accuracy needed for a high-performance (HP) tool to underpin global food security, as well as help advance precision agriculture. Further research can be done to evaluate its scalability on bigger datasets as well as its deployment with IoT based real time monitoring systems to improve its practical implementation.

Future research assessing the crop health analysis system could enhance the existing framework to integrate Internet of Things (IoT)-based real-time monitoring systems like aerial drones and field sensors, allowing continuous data collection and dynamic disease detection. Providing this system more data with multiple crops, and crop diseases in various regions will improve the generalizability of the system. Also, fine-tuning the model for low-resource settings and creating a mobile app with multi-language functionality would enable the technology to reach small-haller farmers, which would encourage widespread usage and maintain sustainable farming habits around the world.

6. Acknowledgement

I would like to express my sincere gratitude to all those who provided support during the completion of this research. While this work was not funded by any external organization or grant, it would not have been possible without the intellectual and moral support of many individuals.

I am deeply thankful to my co-authors (Newaz IbrahimKhan, Noshin Un Noor, Fazle Rabby, Ashiqur Rahman, Hridoy Chandra Shill), for their invaluable guidance, constructive criticism, and unwavering encouragement throughout this research journey. Their expertise and insights significantly shaped this work.

7. References

1. Aggarwal S, Sahoo AK, Bansal C, Sarangi PK. Image Classification using Deep Learning: A Comparative Study of VGG-16, InceptionV3 and EfficientNet B7 Models. In: 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India; c2023. p. 1728-1732. DOI: 10.1109/ICACITE57410.2023.10183255.
2. Ahmed K, Shahidi TR, Irfanul Alam SM, Momen S. Rice Leaf Disease Detection Using Machine Learning Techniques. In: 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh; c2019. p. 1-5. DOI: 10.1109/STI47673.2019.9068096.
3. Amin H, Darwish A, Hassanien AE, Soliman M. End-to-End Deep Learning Model for Corn Leaf Disease Classification. IEEE Access. 2022;10:31103-3115. DOI: 10.1109/ACCESS.2022.3159678.
4. Aminuddin NF, Tukiran Z, Joret A, Tomari R, Morsin M. An Improved Deep Learning Model of Chili Disease Recognition with Small Dataset. International Journal of Advanced Computer Science and Applications. 2022, 13(7). DOI: 10.14569/IJACSA.2022.0130750.
5. Prabin Kumar Sahoo, Deoraj Sharma. Economic impact of artificial intelligence in the field of agriculture. Int J Hortic Food Sci 2023;5(1):29-34. DOI: 10.33545/26631067.2023.v5.i1a.152
6. Ashok S, Kishore G, Rajesh V, Suchitra S, Sophia SGG, Pavithra B, *et al*. Tomato Leaf Disease Detection Using Deep Learning Techniques. In: 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India; c2020. p. 979-983. DOI: 10.1109/ICCES48766.2020.9137986.
7. Shaikh TA, Rasool T, Lone FR. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. Computers and Electronics in Agriculture. 2022;198:107119. DOI: 10.1016/j.compag.2022.107119.
8. B R P, Ashok A, S H A V S. Plant Disease Detection and Classification Using Deep Learning Model. In: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India; c2021. p. 1285-1291. DOI: 10.1109/ICIRCA51532.2021.9544729.
9. Basavaiah J, Anthony A. Tomato Leaf Disease Classification using Multiple Feature Extraction Techniques. Wireless Pers Commun. 2020;115:633-651. DOI: 10.1007/s11277-020-07590-x.
10. Chakravarthy AS, Raman S. Early Blight Identification in Tomato Leaves using Deep Learning. In: 2020 International Conference on Contemporary Computing and Applications (IC3A), Lucknow, India; c2020. p. 154-158. DOI: 10.1109/IC3A48958.2020.233288.
11. Domingues T, Brandão T, Ferreira JC. Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey. Agriculture. 2022;12(9):1350. DOI: 10.3390/agriculture12091350.
12. Husin ZB, Shakaff AYBM, Aziz AHBA, Farook RBSM. Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques. In: 2012 Third International Conference on Intelligent Systems Modelling and Simulation, Kota Kinabalu, Malaysia; c2012. p. 291-296. DOI: 10.1109/ISMS.2012.33.
13. Jafar A, Bibi N, Naqvi RA, Sadeghi-Niaraki A, Jeong D. Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations. Frontiers in Plant Science. 2024, 15. DOI: https://doi.org/10.3389/fpls.2024.1356260.