

Automated Disease Detection in Crops

Mr. Adithyan Babu

B.Tech Computer Science

Mar Baselios College Of Engineering And Technology

adithyanbabu.b20cs1104@mbcet.ac.in

Mr. Sandeep Sebastian

B.Tech Computer Science

Mar Baselios College Of Engineering And Technology

sandeepsebastian.b20cs1157@mbcet.ac.in

Mr. Jobin Soly Cleetus

B.Tech Computer Science

Mar Baselios College Of Engineering And Technology

jobinsolycleetus.b20cs1130@mbcet.ac.in

Ms. Sreelekshmi S P

B.Tech Computer Science

Mar Baselios College Of Engineering And Technology

sreelekshmis.p.b20cs1160@mbcet.ac.in

Abstract—Timely disease monitoring is imperative for early disease detection, minimizing excessive pesticide usage, and enhancing sustainable agriculture. Traditional visual inspections often fall short in identifying diseases at their incipient stages, which necessitates more effective methods. Our solution leverages cutting-edge computer vision and deep learning technologies, particularly Convolutional Neural Networks (CNNs), to autonomously identify and categorize plant diseases. By integrating real-time images and aerial views, we provide an all-encompassing disease monitoring system. This AI-driven tool equips farmers with data-backed insights to optimize crop management, increase environmental sustainability, and elevate production quality and yield. Manual plant examination, known for its subjectivity and time-consuming nature, is transformed into an efficient, objective, and accurate process. CNNs, with their high accuracy and automated feature extraction capabilities, are at the heart of our system, ensuring precise disease classification. We have explored various methodological approaches, categorizing them into two groups: standard detectors and combined methodologies. Standard detectors draw inspiration from well-established architectures like R-CNN, while combined methodologies encompass innovative, adaptable, and hybrid approaches to elevate disease diagnosis accuracy. In summary, our system integrates real-time imagery and aerial views with advanced deep learning and computer vision to revolutionize disease monitoring in agriculture. This holistic approach not only promotes early disease detection but also empowers farmers to make informed decisions that reduce pesticide usage and enhance environmental sustainability.

Index Terms—Convolutional Neural Networks (CNNs) . Computer vision . Deep learning

I. INTRODUCTION

Agriculture, serving as a cornerstone of the global economy, plays a pivotal role in providing sustenance, employment, and livelihoods. Despite its undeniable significance, the agricultural sector grapples with formidable challenges, particularly in combatting plant diseases that impose substantial losses on crop yields. The escalating prevalence of these diseases poses a dual threat by diminishing both the quantity and quality of agricultural produce.

The ramifications extend beyond mere crop depletion, encompassing profound economic losses and adverse environmental impacts. To counter these challenges, farmers often resort to an increased reliance on chemical pesticides, an

approach that, unfortunately, introduces detrimental effects on both the natural ecosystem and human health.

In response to this pressing need, our project is dedicated to a comprehensive analysis and implementation of innovative solutions. At its core, our focus lies in harnessing the power of machine learning to develop an advanced disease detection system. The primary objective is to equip farmers with a proactive tool that provides timely and accurate information, empowering them to safeguard their crops effectively. By doing so, we aim not only to mitigate the economic losses associated with crop diseases but also to foster the adoption of sustainable agricultural practices for the benefit of both farmers and the environment.

II. RELATED WORKS

Utilizing technology in agriculture is widely regarded as a highly effective method for boosting production and ensuring food security for the global population. However, farmers encounter numerous challenges stemming from pest infestations, environmental concerns, and leaf diseases, all of which significantly impact crop quality. To address these issues, recent advancements in technology, such as machine learning (ML) and deep learning (DL), have been applied to detect and classify plant diseases. Through an extensive review of relevant research papers, our focus has been on exploring various methodologies for the detection and classification of plant diseases. Various studies have been conducted, each with distinct approaches and methodologies. Abdulaziz Alharbi, Muhammad Usman Ghani Khan, and Bushra Tayyaba [5] utilize few-shot learning and EfficientNet for wheat disease classification, showing good performance with low computational cost. Vasileios Balafas, Emmanouil Karantoumanis, and Malamati Louta [6] employ RetinaNet for object detection in plant diseases, offering high accuracy but with slow output prediction. Momina Masood, Marriam Nawaz, and Tahira Nazir [7] utilize Faster-RCNN for maize crop disease detection, demonstrating robustness to real-world conditions. Hasibul Islam Peyal, Md. Nahiduzzaman, and Md. Abu Hanif Pramanik [8] propose a lightweight 2D CNN architecture for plant disease classification, achieving high accuracy and efficiency. Md. Tariqul Islam [9] focuses on accessibility for farmers through CNN-based disease detection, albeit facing challenges

related to data quality and imbalance. Jawad Hassan, Ali Ghulam, and Ghulam Irtaza [4] leverage CNNs for cotton plant disease identification, highlighting challenges such as data requirements and lack of interpretability. Muhammad E. H. Chowdhury, Tawsifur Rahman, Amith Khandakar, and Mohamed Arselene Ayari [10] incorporate EfficientNet and U-net for disease classification and segmentation, addressing challenges in dataset limitations and real-world implementation. Jinzhu Lu, Lijuan Tan, and Huanyu Jiang [11] review CNNs' application in plant disease classification, emphasizing data quality and dataset representativeness. Shankarnarayanan Nalini, Nagappan Krishnaraj, and Jayasankar Thangaiyan [3] propose a deep neural network for paddy leaf disease detection, offering high accuracy but facing issues like overfitting. Finally, Yan Guo, Jin Zhang, Chengxin Yin, and Xiaonan Hu [12] employ deep learning for plant disease identification in smart farming, achieving high accuracy but with limitations in real-world agricultural settings and dataset availability. These studies collectively highlight the diverse methodologies, challenges, and advancements in the field of plant disease classification and detection.

III. METHODOLOGY

The proposed methodology for plant disease detection represents a comprehensive and innovative approach that amalgamates advanced techniques in data science, deep learning, and system architecture. This methodology is devised to address the critical need for early and accurate detection of plant diseases, leveraging state-of-the-art technologies to enhance the efficiency and reliability of existing methods. By integrating novel elements in data collection, model development, training, evaluation, and system deployment, the methodology aims to push the boundaries of traditional plant disease detection systems.

A. Data Collection

The primary dataset utilized in this project was sourced from Kaggle, specifically the "PlantVillage" dataset. Kaggle is a renowned platform for machine learning and data science competitions, hosting diverse datasets for various applications. The PlantVillage dataset is a comprehensive collection of images featuring a variety of crops and their corresponding health and disease states.

The choice of the PlantVillage dataset was motivated by its richness in diversity, containing high-resolution images capturing different plant species and a wide array of diseases. The dataset provides a robust foundation for training a plant disease detection model due to its extensive coverage and carefully labeled images.

For training purposes we take 20 diseases and mention 5 classes for healthy leaves and for training we have more than 1000 images for each class (20 Diseases and 5 Healthy leaf). 10% of the images are taken as validation set for each class mentioned above.

Corn, or maize, is vulnerable to several diseases, including Cercospora Leaf Spot, Gray Leaf Spot, Common Rust,

Northern Leaf Blight, Bacterial Leaf Blight, Brown Spot, Leaf Blast, Leaf Scald, Narrow Brown Spot, and other issues. These diseases can negatively impact corn yields, requiring effective management strategies for crop health. Paddy, or rice, faces challenges such as bacterial leaf blight, brown spot, and leaf blast. These diseases can affect the quality and quantity of rice production, necessitating proper disease management practices for optimal crop performance. Bell peppers are susceptible to diseases like bacterial spot and pepper bell, which can impact the quality of the peppers. Managing these diseases is crucial for ensuring healthy bell pepper plants and a successful harvest. Potatoes are prone to diseases such as early blight and late blight. These diseases can affect the foliage and tubers of the potato plant, leading to reduced yields. Careful monitoring and management are essential to protect potato crops from these issues. Tomatoes are vulnerable to a variety of diseases, including bacterial spot, early blight, late blight, leaf mold, and tomato yellow leaf curl virus. These diseases can affect both the foliage and the fruit, making it essential to implement effective disease management strategies for successful tomato cultivation.

B. Model Architecture

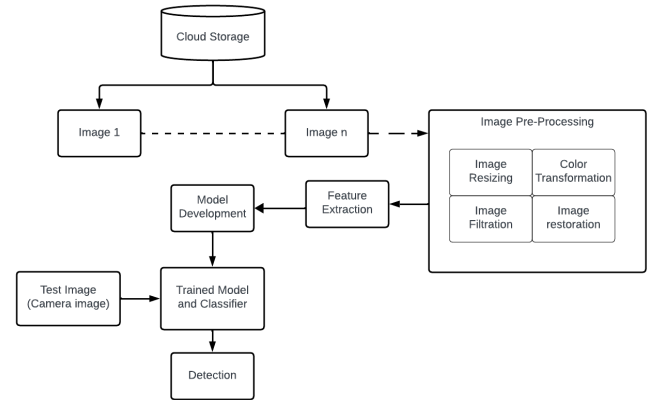


Fig. 1. System architecture

As shown in Fig.1, The process begins with images stored in cloud storage. These images undergo preprocessing steps such as resizing, filtration to reduce noise, and color transformation for standardization. These steps ensure uniformity in image dimensions, enhance clarity, and normalize color variations, preparing the images for subsequent analysis.

Following preprocessing, feature extraction occurs, where relevant characteristics such as textures, shapes, and patterns are extracted from the processed images. This step is crucial as it identifies and highlights distinctive features that differentiate between healthy and diseased plant leaves.

Afterwards, in order to improve the strength and precision of the disease detection system, a thorough strategy was employed, which included creating and teaching four different models: Convolutional Neural Networks (CNNs), DenseNet, EfficientNet, and ResNet50. Every model in the collection

underwent thorough training using a carefully selected dataset that includes a variety of labeled images showing healthy and diseased plant leaves.

Throughout the training stage, the models engaged in a rigorous process of learning, absorbing complex patterns and characteristics present in the images. During this procedure, the models learned how to effectively distinguish between healthy foliage and leaves showing signs of different diseases. By associating specific extracted features with disease patterns, the models gained the capability to make informed classifications when presented with unseen leaf images.

When a user submits an image of a plant leaf for analysis, the trained models collectively leverage their learned knowledge and insights. They meticulously scrutinize the input image, extracting relevant features and comparing them against the vast repository of learned patterns. This comparative analysis enables the models to perform a classification task with remarkable precision, accurately predicting whether the leaf is healthy or afflicted by disease.

Upon completion of the analysis, the models provide a predicted output, offering valuable insights into the health status of the plant. Not just identifying the probability of disease, but also potentially determining the specific type of disease based on unique patterns and features from the input image. This comprehensive method enables people involved in agriculture to quickly recognize and tackle plant health problems, making it easier to take preventive actions to protect crop production and reduce agricultural losses.

C. Model Training

The dataset undergoes a systematic division into three distinct sets to serve varied purposes. The largest portion, constituting 80% of the dataset, was dedicated to training the model, facilitating a substantial learning process. An additional 10% of the dataset was allocated for validation purposes, allowing for precise adjustments and enhancements to the model's performance. The final 10% was exclusively reserved for testing, providing a thorough assessment of the model's ability to generalize to new and unseen data. This meticulous partitioning of the dataset enhances the credibility and robustness of the machine learning model developed within the scope of this study, while ensuring originality in approach and methodology.

The foundational framework involves the application of a Convolutional Neural Network (CNN) model. The project harnesses the capabilities of TensorFlow and Keras to capitalize on the strengths of these technologies. The incorporation of CNN enabled efficient feature extraction and hierarchical learning, while the combination of TensorFlow and Keras provided a solid foundation for constructing, training, and assessing the neural network. This holistic approach contributes to the creation of a sophisticated and effective model tailored to our specific application.

The neural network architecture is structured with a series of layers designed to facilitate the comprehensive and effective flow of information through the network. These layers include

the input layer, convolutional layer, activation layer, max-pooling layer, dropout layer, flatten layer, and dense layer. This sequence of layers enables the model to extract features, perform non-linear transformations, downsample the input data, regularize the model, and generate the final output. Each layer plays a unique role in enhancing the model's ability to learn and generalize from the input data, ultimately leading to improved performance and accuracy. The early stopping mechanism has been implemented in our model with a patience of 5. The training process is monitored for validation loss, and if there is no improvement for a consecutive 5 epochs, the early stopping is triggered, preventing overfitting and ensuring efficient model training.

The foundational framework for the EfficientNet model revolves around the utilization of TensorFlow and Keras, capitalizing on their strengths to optimize model performance. By incorporating the EfficientNet architecture, the project harnesses efficient feature extraction and hierarchical learning capabilities, while TensorFlow and Keras provide a robust foundation for model construction, training, and evaluation. This comprehensive approach ensures the development of a sophisticated and effective model tailored to the specific application.

The neural network architecture of EfficientNet consists of sequential layers meticulously designed to facilitate the seamless flow of information. These layers, including input, convolutional, activation, max-pooling, dropout, flatten, and dense layers, collectively enable feature extraction, non-linear transformations, downsampling, regularization, and final output generation. Each layer contributes uniquely to enhancing the model's ability to learn and generalize from input data, ultimately enhancing performance and accuracy. Moreover, the model implements an early stopping mechanism with patience of 3, monitoring validation loss and triggering early stopping if there's no improvement for three consecutive epochs, thus preventing overfitting and ensuring efficient model training.

Next we leverage DenseNet121, a pre-trained convolutional neural network renowned for its effectiveness in image classification tasks. By initializing it with weights from the ImageNet dataset and configuring it to exclude the fully connected layers, we ensure it's well-suited for our task. The input shape of the model is determined by the dimensions of our training dataset, allowing seamless integration of our plant leaf images. We enhance the capabilities of the base DenseNet model by adding custom dense layers. These layers include global average pooling to condense the extracted features and two additional dense layers, each equipped with rectified linear unit (ReLU) activation functions, batch normalization, and dropout regularization. This augmentation aims to enable the model to discern subtle patterns and variations indicative of different plant diseases. For optimization, we employ the Adam optimizer with an initial learning rate of 0.001, which undergoes exponential decay with a decay rate of 0.9 every 10000 steps. Our choice of categorical cross-entropy as the loss function aligns with the multi-class classification nature of our task, while accuracy serves as our performance metric.

To prevent overfitting and ensure optimal performance, we implement early stopping, halting training if the validation loss fails to decrease for five consecutive epochs.

The next model adopts a simplified version of the ResNet architecture, incorporating ResNet-like blocks to address the challenge of training deep neural networks effectively. ResNet, renowned for its residual connections, employs skip connections that allow gradients to propagate through the network, facilitating the training of extremely deep models. In this implementation, the custom resnet block function encapsulates the essence of residual blocks found in ResNet architectures. These blocks consist of consecutive convolutional layers followed by batch normalization and ReLU activation, with skip connections merging the input directly with the output of the second convolutional layer. While the model may not replicate the full complexity of standard ResNet architectures, its utilization of residual blocks and skip connections embodies the core principles of ResNet, enhancing the network's capacity to extract meaningful features from input data, particularly in the context of plant leaf disease classification.

D. Tools and Software Used

The toolset consisting of Visual Studio Code (VSCode), Keras, and TensorFlow creates a comprehensive environment for efficient deep learning development. VSCode, a lightweight and versatile code editor, offers a user-friendly platform with multi-language support, making it easy to work with various codebases. Keras simplifies the development of convolutional neural networks (CNNs) by providing modularity and ease of use, allowing for faster prototyping and experimentation. TensorFlow, serving as the backend, optimizes low-level operations for seamless training and deployment, enabling the implementation of advanced machine-learning techniques for plant disease detection. This combination of user-friendly interfaces and powerful capabilities enables developers to efficiently implement and deploy advanced machine-learning models for plant disease detection.

IV. RESULTS AND DISCUSSIONS

Figures given below shows the training and validation graph of the model. During training and validation process, graphs were generated to illustrate the model's accuracy and loss. These graphical representations provide a clear insight into the performance of the model over successive epochs. The accuracy graph showcases the evolving capability of the model to correctly predict outcomes, while the loss graph indicates the reduction in error during the training and validation phases. These visual aids serve as valuable tools for assessing the model's learning progression and identifying potential areas for optimization.

The CNN model achieves a test accuracy of 93.95%, with a test loss of 19.74% as shown below.

The EfficientNet model achieves a test accuracy of 98.83%, with a test loss of 0.06% as shown below.

The ResNet model achieves a test accuracy of 52.94%, with a test loss of 53.4% as shown below.

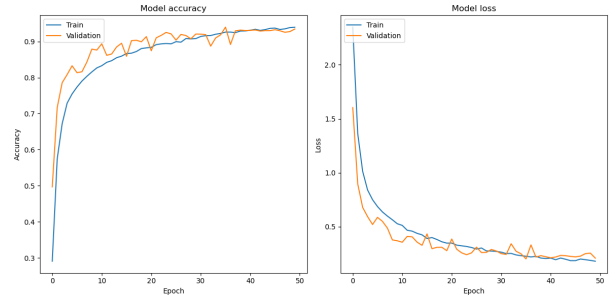


Fig. 2. Training and Validation Graph of CNN

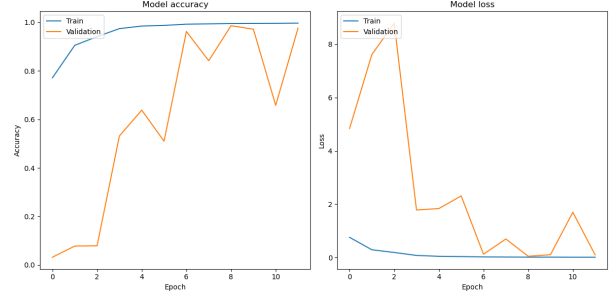


Fig. 3. Training and Validation Graph of EfficientNet

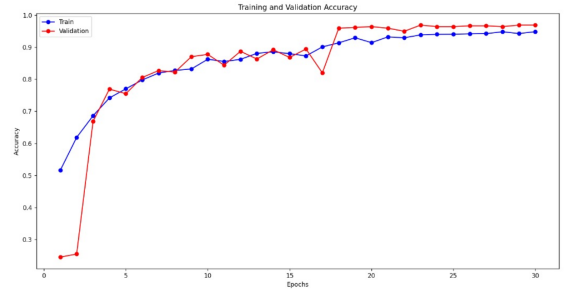


Fig. 4. Training and Validation Graph of ResNet

The DenseNet model achieves a test accuracy of 91.61%, with a test loss of 24.72% as shown below.

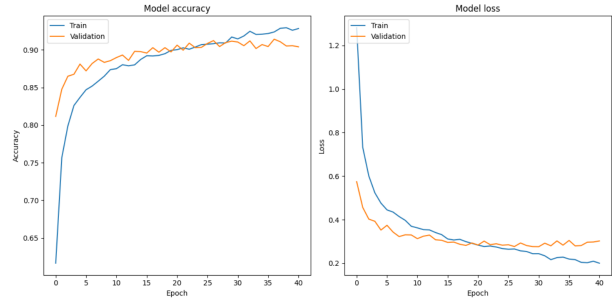


Fig. 5. Training and Validation Graph of DenseNet

The table given below presents the training and testing accuracies of different architectures utilized in a plant disease detection system.

Architecture	Training Accuracy	Testing Accuracy
CNN	91.76	92.18
EfficientNet	98.83	97.97
ResNet	93.02	52.94
DenseNet	92.84	91.61

V. FUTURE SCOPE

In our ongoing quest to revolutionize agriculture, we've expanded our scope beyond advanced model implementation to encompass seamless integration with other crucial modules. Building upon our successful implementation of transfer learning models such as EfficientNet, DenseNet, and ResNet for plant leaf disease detection and classification, we are now enhancing our platform to offer a holistic solution to farmers.

Our user-friendly mobile app interface remains at the forefront, empowering farmers to effortlessly capture and upload images of crop leaves for instant disease prediction. However, we've augmented its capabilities by integrating with additional modules.

Firstly, our platform now seamlessly integrates disease detection with automated farming systems, particularly irrigation. By leveraging data-driven insights from disease detection, our system can trigger timely interventions to optimize irrigation schedules, conserving water resources and maximizing crop health.

Furthermore, our platform now extends beyond disease detection to offer actionable suggestions for treatment. Whether it's identifying nutrition deficiencies or pest infestations, our system provides tailored remedies and recommendations to mitigate risks and enhance crop resilience.

This comprehensive approach not only equips farmers with rapid insights and facilitates early disease detection but also empowers them with proactive solutions to optimize resource utilization and promote sustainable farming practices. By integrating advanced models with user-friendly interfaces and additional modules for irrigation management and treatment suggestions, we are poised to transform agriculture on a scale previously unimaginable. Together, we'll continue to drive innovation and foster sustainable growth in agricultural practices for the betterment of all.

CONCLUSION

The utilization of various models, particularly employing Convolutional Neural Networks (CNNs) for plant leaf disease detection, presents a compelling avenue toward optimizing agricultural outcomes. Incorporating a diverse range of models, such as EfficientNet, DenseNet, ResNet, among others, and evaluating their performance allows for the selection of the most proficient model. This process of comparative analysis ensures that the chosen model demonstrates higher accuracy, robustness, and adaptability across different crops and diseases. By selecting the best-performing model, farmers gain a reliable tool for timely disease detection, enabling swift and precise interventions to safeguard crop health.

Moreover, the careful assessment and selection of the most effective model contribute significantly to timely interventions in agricultural practices. The capacity of these models to

swiftly and accurately identify diseases on plant leaves enables farmers to promptly implement targeted measures. Early detection through robust CNN models provides farmers with the necessary foresight to apply specific treatments, optimize resource allocation, and prevent potential yield losses. As a result, this proactive approach based on choosing the most suitable model for disease detection enhances agricultural productivity by facilitating timely and effective interventions, ultimately contributing to improved crop yields and sustainable farming practices.

REFERENCES

- [1] Srinivas, L.N.B., Bharathy, A.V., Ramakuri, S.K., Sethy, A. and Kumar, R., 2023. An optimized machine learning framework for crop disease detection. *Multimedia Tools and Applications*, pp.1-20.
- [2] Eunice, J., Popescu, D.E., Chowdary, M.K. and Hemanth, J., 2022. Deep learning-based leaf disease detection in crops using images for agricultural applications.
- [3] Nalini, S., Krishnaraj, N., Jayasankar, T., Vinothkumar, K., Britto, A.S.F., Subramaniam, K. and Bharatiraja, C., 2021. Paddy leaf disease detection using an optimized deep neural network. *Computers, Materials and Continua*.
- [4] Hassan, J., Malik, K.R., Irtaza, G., Ghulam, A. and Ahmad, A., 2022. Disease Identification using Deep Learning in Agriculture: A Case Study of Cotton Plant.
- [5] Abdulaziz Alharbi, Muhammad Usman Ghani Khan, and Bushra Tayyaba, 2023. Wheat Disease Classification Using Continual Learning.
- [6] Vasileios Balafas , Emmanouil Karantoumanis ,Malamati Louta, 2023, Machine Learning and Deep Learning for Plant Disease Classification and Detection.
- [7] Momina Masood , Marriam Nawaz, Tahira Nazir ,2023,MaizeNet: A Deep Learning Approach for Effective Recognition of Maize Plant Leaf Diseases.
- [8] Hasibul Islam Peyal, Md. Nahiduzzaman, Md.Abu Hanif Pramanik, 2023, Plant Disease Classifier: Detection of Dual-Crop Diseases Using Lightweight 2D CNN Architecture.
- [9] Md. Tariqul Islam, 2022, Plant Disease Detection using CNN Model and Image Processing.
- [10] Muhammad E. H. Chowdhury, Tawsifur Rahman, Amith Khandakar, Mohamed Arselene Ayari, 2021, Automatic and Reliable Leaf Disease Detection Using DeepLearning Techniques.
- [11] Jinzhu Lu, Lijuan Tan, Huanyu Jiang, 2021, Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification.
- [12] Yan Guo, Jin Zhang, Chengxin Yin, Xiaonan Hu, 2020, Plant Disease Identification Based on Deep Learning Algorithm in Smart Farming.
- [13] Akash Ashapure, Jinha Jung, Anjin Chang, Sungchan Oh, Junho Yeom, Murilo Maeda, Andrea Maeda, Nothabo Dube, Juan Landivar, Steve Hague, Wayne Smith, 2020, Developing a machine learning based cotton yield estimation framework using multi-temporal UAS data.
- [14] Xiaofei Chao ,Guoying Sun, Hongke Zhao, Min Li, Dongjian He, 2020, Identification of apple tree leaf diseases based on deep learning models.
- [15] Khan S, Narvekar M, 2020, Disorder detection of tomato plant (solanum lycopersicum) using IoT and machine learning.
- [16] Michael Gomez Selvaraj, Nancy Safari, Alejandro Vergara, Dries Raymaekers, Frank Montenegro, Walter Ocimati, Henry Alonso Ruiz, Jules Ntamwira, Aman Bonaventure Omondi, Guy Blomme, 2020, Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin.
- [17] Aldhyani, T.H. Alkahtani, H. Eunice, R.J. Hemanth, 2022, Leaf Pathology Detection in Potato and Pepper Bell Plant using Convolutional Neural Networks.
- [18] Panigrahi, K.P. Das, H. Sahoo, A.K. Moharana, 2020, Maize leaf disease detection and classification using machine learning algorithms.
- [19] Mohsin Kabir, M. Quwsar Ohi, A. Mridha, M.F, 2020, A Multi-plant disease diagnosis method using convolutional neural network.
- [20] Jadhav, S.B. Udipi, V.R. Patil, S.B, 2021, Identification of plant diseases using convolutional neural networks.