Abstract

Agriculture, a cornerstone of civilization, holds significance on par with technological advancements. One of the primary challenges in agriculture is safeguarding plants against diseases caused by insects and natural adversaries. Historically, farmers have relied on visual inspection to assess the health of plants. However, contemporary technological and Data Science advancements offer innovative methods to determine plant health. In this project, Convolutional Neural Network (CNN) emerges as a valuable tool for disease detection in plants. By capturing images of plant leaves and inputting them into a trained model, CNN facilitates the identification of specific diseases. The utilization of CNN in plant disease detection has demonstrated an accuracy rate of 86%, showcasing its effectiveness. While alternative methods exist, the aim is to streamline and simplify the process of identifying and addressing plant diseases.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction:

Agricultural production, an age-old method of obtaining food, serves as a crucial source of income worldwide. Essential for both humans and animals, plants provide food, oxygen, and other necessities. Governments and experts globally are actively working to enhance food production, yielding success in the real world. When plants succumb to diseases, it affects the entire environment, impacting living organisms. Plant diseases can manifest in various parts, such as stems, leaves, and branches, with diverse types like bacterial and fungal diseases influenced by factors like climate.

Food insecurity affects many due to insufficient crop output, exacerbated by climate changes affecting plant development. Early detection of plant diseases is vital for preventing large-scale crop losses. Farmers need to apply appropriate insecticides, avoiding excessive use harmful to crops and farmland. Seeking expert advice helps in proper chemical application. Researchers focus on plants to assist those in agriculture. Visible diseases are easily detectable, but early detection relies on consistent monitoring.

Innovations such as automated disease detection tools benefit farmers in both small and large-scale cultivation. These tools provide accurate results, aiding in the precise identification and management of plant disorders.

Cutting-edge technologies have enabled the rapid detection of plant diseases, leveraging the power of deep learning and neural networks. In this study, a Deep Convolutional Neural Network (CNN) plays a pivotal role in identifying both infected and healthy leaves, providing a means to detect illnesses in afflicted plants. The CNN model is tailored to accommodate variations between healthy and diseased leaves. Training the model involves using images of leaves, and the output is determined based on the characteristics of the input leaf. This approach ensures a swift and accurate identification of plant health status.

Fig 1.1 : Sample Images from the dataset (from left to right) Rust, Healthy

Figure 1.1 shows Rust and healthy leaves represent two distinct states in the visual diagnosis of lant health. In the context of plants, rust is a common fungal disease characterized by reddish-brown discoloration on the leaves, often resembling the appearance of rusted metal. This condition is caused by various fungi that thrive in humid environments, affecting the overall vitality of the plant. On the contrary, healthy leaves exhibit a vibrant and uniform green color, showcasing the plant's well-being. In a diagnostic diagram, rust-affected leaves may feature irregular patches or lesions, indicative of the fungal infection, while healthy leaves would display a consistent and undamaged surface. The juxtaposition of these visual cues in a diagram serves as a valuable tool for farmers and researchers to quickly identify and differentiate between healthy and rust-infected leaves, facilitating timely intervention and management strategies to protect plant health and optimize agricultural yield.

Plant disease detection is a critical aspect of modern agriculture, where the timely identification of ailments affecting crops can significantly impact yield and food security. With advancements in technology, particularly in the field of computer vision and deep learning, innovative approaches have emerged for efficient and accurate detection of plant diseases. Among these, Convolutional Neural Networks (CNN) and Sequential models play a pivotal role in automating the process, providing a robust solution to the challenges faced by farmers and researchers in monitoring and managing crop health. Convolutional Neural Networks (CNNs) have proven to be highly effective in image recognition tasks, making them particularly well-suited for plant disease detection. These neural networks are adept at learning intricate patterns and features within images, enabling them to discriminate between healthy and diseased plant tissues. By training CNN models on diverse datasets of plant images, encompassing both healthy and infected samples, the network learns to recognize subtle visual cues indicative of diseases.

Sequential models, often employed in conjunction with CNNs, facilitate the development of comprehensive frameworks for disease detection. These models allow for the sequential arrangement of layers, enabling the integration of convolutional, pooling, and dense layers for feature extraction, abstraction, and classification. The sequential architecture provides a structured framework for building sophisticated models capable of capturing intricate relationships within the data.

The combination of CNNs and Sequential models in plant disease detection offers a scalable and efficient solution. By harnessing the power of deep learning, these models contribute to early and accurate identification of diseases, empowering farmers with t he information needed to implement timely interventions. This not only enhances crop yield but also aids in the sustainable management of agricultural resources. The integration of these technologies marks a significant stride towards precision agriculture, where data-driven insights contribute to more informed and proactive decision-making in the realm of plant health management.

Chapter 2

NOVELTY/CONTRIBUTION

Our project introduces a novel approach by incorporating advanced deep learning techniques, specifically Convolutional Neural Networks (CNN) and Sequential models. While previous methodologies may have relied on traditional image processing or basic machine learning algorithms, our contribution lies in the utilization of state-of-the-art CNNs for their unparalleled ability to automatically learn complex patterns and features from plant images. Additionally, the integration of Sequential models provides a structured framework for more sophisticated data processing, allowing for comprehensive feature extraction and classification. This departure from conventional methods represents a substantial leap forward in the accuracy and efficiency of plant disease detection. By embracing these advanced neural network architectures, our project aims to significantly enhance the precision, scalability, and automation of the detection process, thereby contributing to the evolution of more robust and effective solutions for monitoring and managing plant health in agriculture.

Chapter 3

DATASET DECSCRIPTION: KAGGLE

The dataset is taken from Kaggle which is structured into distinct training, test, and validation subsets, each containing three subfolders corresponding to the different leaf conditions: healthy, powdery, and rusty. In the training set, there is a considerable diversity with 458 images of healthy leaves, 430 images of powdery leaves, and 434 images of rusty leaves. This robust representation across classes ensures a comprehensive training regime for the machine learning model. Moving to the test set, there are 50 images per class, totaling 150 images, while the validation set consists of 20 images per class, totaling 60 images. This meticulous organization of the dataset provides a balanced and varied collection of plant leaf images for training, rigorous testing, and model validation. The aim is to equip the model with the capability to accurately classify and detect different leaf conditions, ultimately contributing to enhanced precision in plant disease identification and management.

Chapter 4

DESIGN

To develop a Convolutional Neural Network (CNN) for disease prediction, the first step involves loading the dataset. Subsequently, the images are converted into NumPy arrays to facilitate numerical operations. The dataset is then split into training and test sets for model evaluation. The CNN model is built, consisting of convolutional layers for feature extraction and dense layers for classification. Training the model involves feeding it with the training set and validating its performance on the test set. The objective is to predict diseases based on image features, and the CNN is fine-tuned iteratively for optimal predictive accuracy during the training process.

In the realm of plant disease detection using Convolutional Neural Networks (CNN) with Keras and TensorFlow, the process begins by loading a diverse dataset comprising images of healthy leaves and leaves afflicted with either a powdery or rusty disease. These images are then converted into NumPy arrays, preparing them for input into the CNN model. Subsequently, the dataset is meticulously split into training, test, and validation sets to ensure robust model training and evaluation. The CNN model, defined using the Sequential API in Keras, incorporates convolutional layers for feature extraction, max-pooling layers for spatial reduction, and dense layers for classification. The architecture aims to discern patterns indicative of various diseases. With 458 healthy, 430 powdery, and 434 rusty leaf images in the training set, the model undergoes training and validation, exhibiting its prowess in distinguishing between healthy and diseased plants. The test set, with 50 images for each category, serves as a reliable benchmark for assessing the model's predictive accuracy in real-world scenarios. This comprehensive approach underscores the effectiveness of CNNs in plant disease detection, paving the way for enhanced crop management practices.

4.1 Flow chart

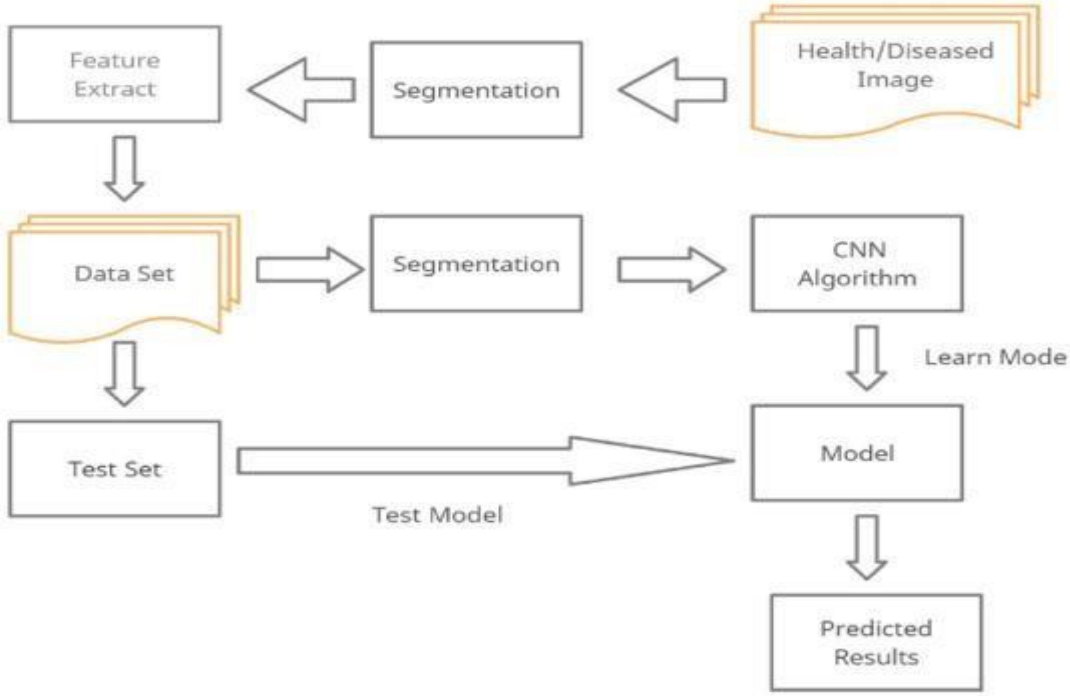

Fig 4.1: workflow of proposed system

In the domain of plant disease detection utilizing Convolutional Neural Networks (CNN) through the Sequential API in Keras and TensorFlow, the process unfolds through a structured flowchart. Initially, a dataset comprising images of both healthy and diseased leaves is employed. The images undergo segmentation to distinguish key regions of interest, and subsequent feature extraction enables the model to focus on pertinent aspects for classification. This curated dataset is then used to train a CNN algorithm, employing the Sequential model architecture in Keras. The CNN learns intricate patterns and relationships within the segmented images, enhancing its ability to discern between healthy and diseased leaves. The model is rigorously tested using a dedicated test set, evaluating its proficiency in predicting disease conditions based on unseen data. The results obtained post-testing provide valuable insights into the model's accuracy and efficacy, highlighting the potential of CNNs in revolutionizing plant disease detection for improved agricultural practices.

Chapter 5

METHODOLOGY

Algorithm: CNN

Input:

- Input image tensor
- Number of classes (for classification tasks)
- Hyperparameters (e.g., learning rate, batch size, etc.)

Output:

Predicted class probabilities (for classification tasks)

Initialization:

1. Define CNN architecture (e.g., convolutional layers, pooling layers, fully connected layers).

2. Initialize network weights (e.g., using random initialization or pre-trained weights).

Training Procedure:

1 Iterate over training batches:
a. Forward pass:

i. Perform basic preprocessing on input images (e.g., normalization).
ii. Pass the preprocessed images through convolutional layers.
iii. Apply activation functions (e.g., ReLU) after each convolutional layer.
iv. Utilize pooling layers to down sample spatial dimensions.
v. Flatten the feature maps.
vi. Pass through fully connected layers.
vii. Apply SoftMax activation on the output layer for multi-class classification.

b. Calculate loss:
i. Compute the SoftMax cross-entropy loss between predicted class probabilities and ground truth labels.

c. Backpropagate gradients:
i. Update network weights using gradient descent optimization algorithms (e.g., Adam, SGD).
2 Repeat training for multiple epochs or until convergence.

Prediction Procedure:

Forward pass:
a. Perform basic preprocessing on input images.
b. Pass preprocessed images through the trained CNN architecture.
c. Compute class probabilities using SoftMax activation on the output layer.

Return predicted class probabilities.

Explanation:

1. Initialization : The raw image data that the CNN will process and the total number of categories or classes in the classification task. Hyperparameters Parameters such as the learning rate, batch size, etc., which control the training process.

2. Training Procedure:

In a Convolutional Neural Network (CNN) designed for image classification, the input comprises raw image data in tensor form, along with information about the number of classes and various hyperparameters such as the learning rate and batch size. The initialization phase involves defining the CNN architecture, specifying the arrangement of layers like convolutional layers, pooling layers, and fully connected layers, and initializing the network weights, either randomly or through pre-trained weights. During the training procedure, the network iterates over batches of training data. In the forward pass, the input images undergo preprocessing, convolutional operations, activation functions (commonly ReLU), pooling layers for spatial down-sampling, flattening of feature maps, and passage through fully connected layers with a final SoftMax activation for multi-class classification. The loss is computed using SoftMax cross-entropy to measure the disparity between predicted class probabilities and actual labels. Gradients are then backpropagated to update the network weights using optimization algorithms such as Adam or SGD. This training process is repeated for multiple epochs until convergence. For predictions, new input images undergo similar preprocessing and are passed through the trained CNN architecture, producing class probabilities using SoftMax activation. The final output consists of the predicted class probabilities for the given input.

CNN Architecture

Convolutional Neural Networks (CNN, or ConvNet) are a type of multi-layer neural network that is meant to discern visual patterns from pixel images. In CNN, ‘convolution’ is referred to as the mathematical function. It’s a type of linear operation in which you can multiply two functions to create a third function that expresses how one function’s shape can be changed by the other.

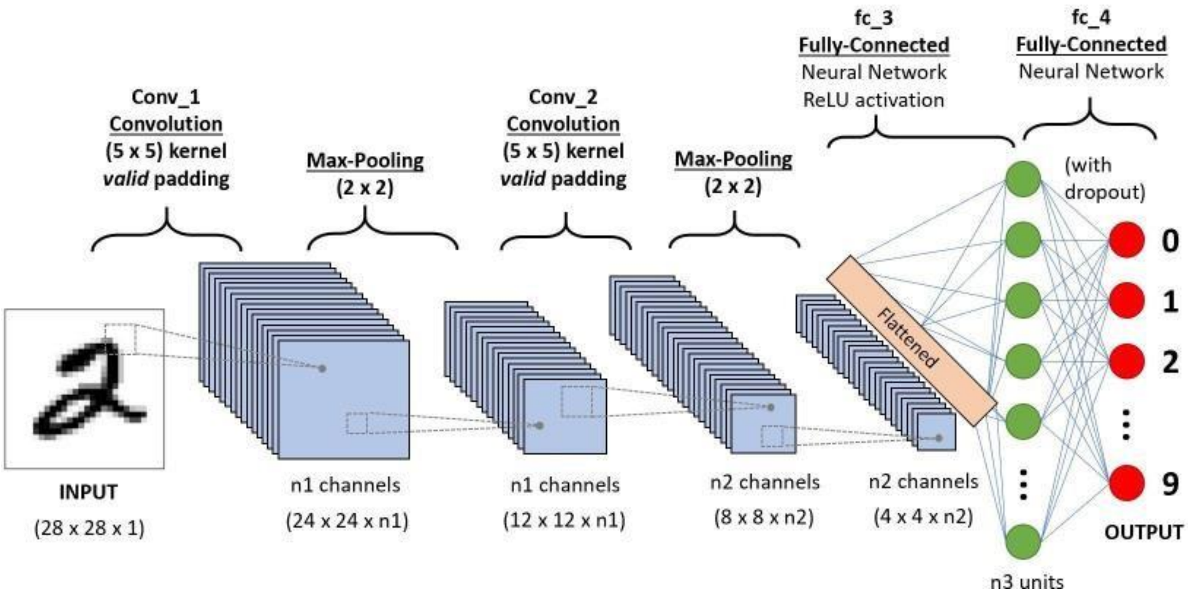

Fig 5.1: Architecture of CNN model

A Convolutional Neural Network (CNN) architecture is a specialized deep learning framework designed for image recognition and computer vision tasks. It comprises several layers, each serving a distinct purpose in the feature extraction process as shown in fig 1.3 . The input layer takes raw image data, and subsequent convolutional layers apply filters to capture local features and patterns. Activation functions like ReLU introduce non-linearity, enhancing the network's capacity to learn complex relationships. Pooling layers reduce spatial dimensions, and flattening transforms the feature maps into a one-dimensional vector. Fully connected layers then connect all nodes, allowing the network to learn global patterns. The output layer produces the final predictions, often using a SoftMax activation for classification tasks.

Various architectures, such as LeNet, AlexNet, and ResNet, differ in depth and design, offering different trade-offs between complexity and performance. Transfer learning, leveraging pre-trained models, is common in CNNs, enabling effective adaptation to specific tasks with limited labeled data. Overall, the intricate design of CNN architectures facilitates automated learning of hierarchical features, making them powerful tools for image-related applications.

Convolution Layer

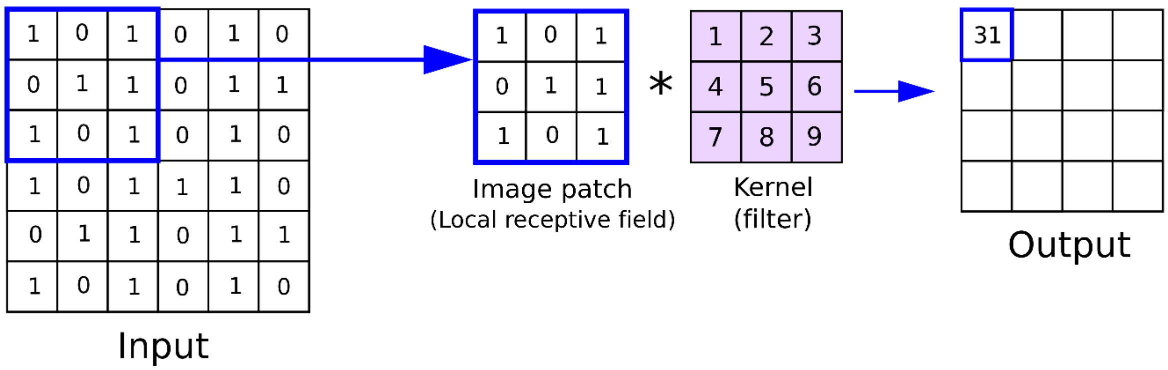

Fig 5.2 : Convolutional layer

The convolutional layer plays a crucial role in Convolutional Neural Networks (CNNs) by generating activation maps through a process of scanning input images using filters. Fig 1.4 illustrates the internal workings of the convolutional layer. In this process, a filter, also known as a kernel, is slid across the input image in a systematic manner, scanning over several pixels at a time. At each position, the filter performs a convolution operation, which involves element-wise multiplication of the filter weights with the corresponding pixel values in the input image, followed by the summation of the results. This process is repeated across the entire image, producing a feature map or activation map that highlights spatial patterns and local features. The learned filters act as feature detectors, capturing important visual information such as edges, textures, or shapes. The convolutional layer's ability to automatically learn and extract relevant features is fundamental to the success of CNNs in image recognition tasks.

Chapter 6

RESULTS

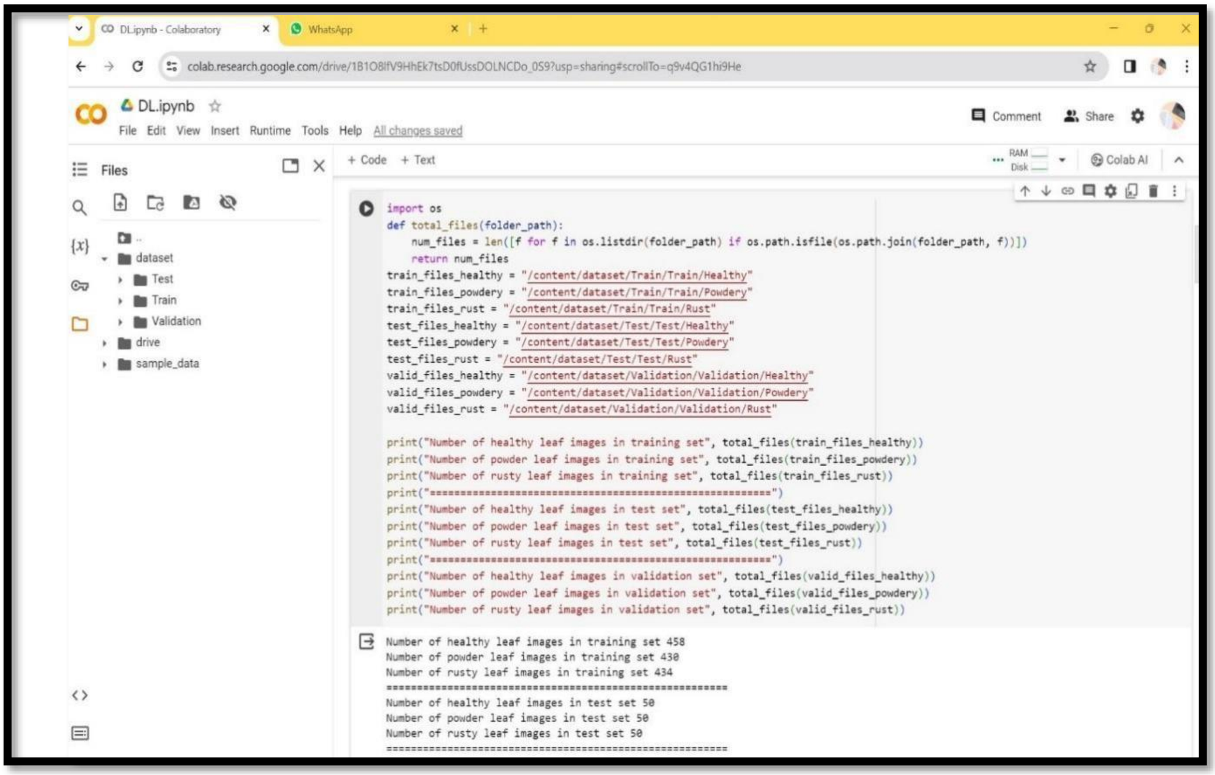

Fig 6.1: Counts of healthy, powdery, and rusty leaf images in training, test, and validation sets.

The provided Python script uses the total_files function to count the number of images in separate directories representing healthy, powdery, and rusty leaves within training, test, and validation sets of a plant leaf dataset. The counts for each category in each set are printed as output shown in figure 6.1. This script is likely a component of a machine learning project, focusing on data distribution analysis for image classification.

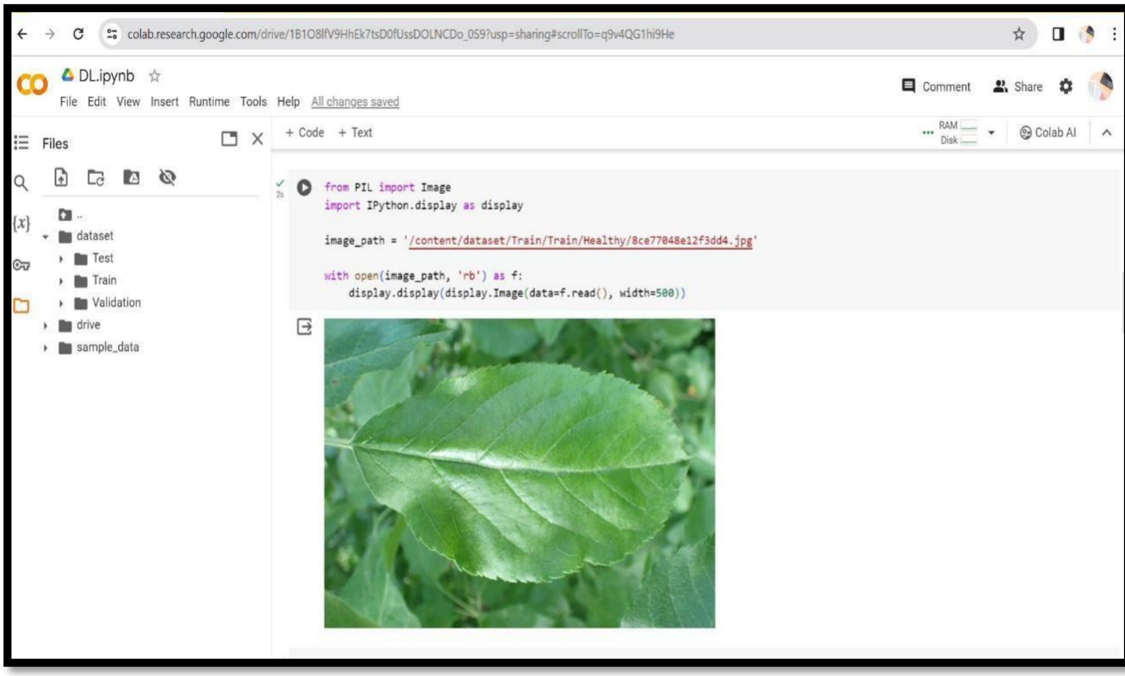

Fig 6.2: Image of healthy leaf in the 'Healthy' category of the dataset.

Figure 6.2 shows part of the 'Healthy' category in the training set, showcases a typical healthy plant leaf. Its visual attributes contribute to training machine learning models to distinguish healthy leaves, marking a vital step in data exploration for image-based machine learning.

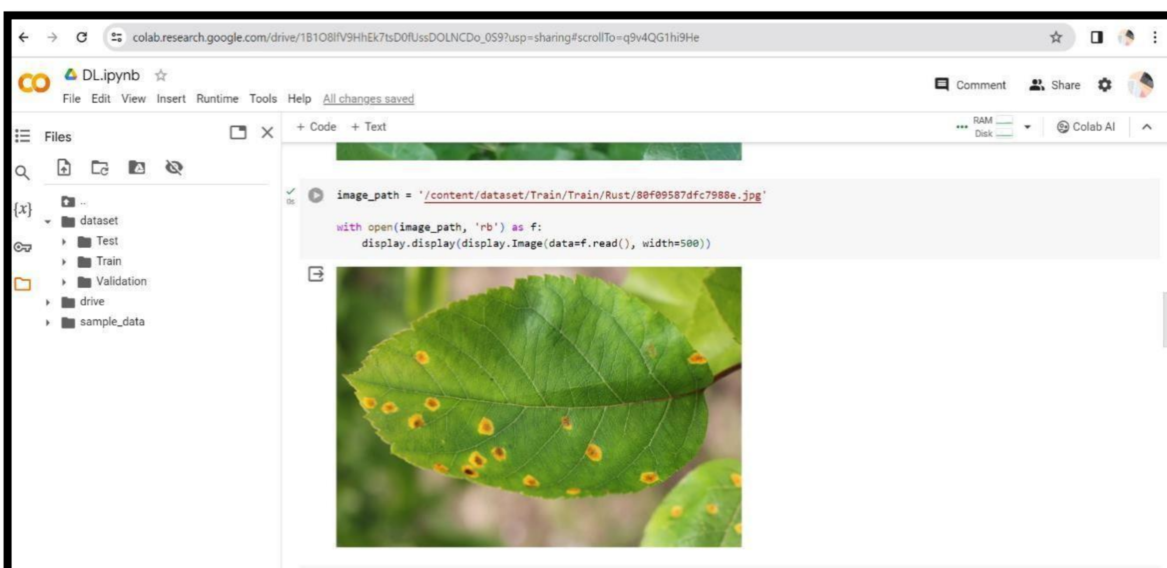

Fig 6.3: Image of Rust leaf in the 'Rust' category of the dataset.

The displayed image, sourced from the 'Rust' category in the training set, depicts a plant leaf affected by rust infection. Visual cues, such as discoloration and lesions, represent characteristic symptoms. Figure 6.3 shows machine learning models to distinguish leaves with rust symptoms, contributing to the dataset's diverse representation of plant health conditions. The successful display of this rust-affected leaf in the output ensures accurate image path validation for model training and evaluation.

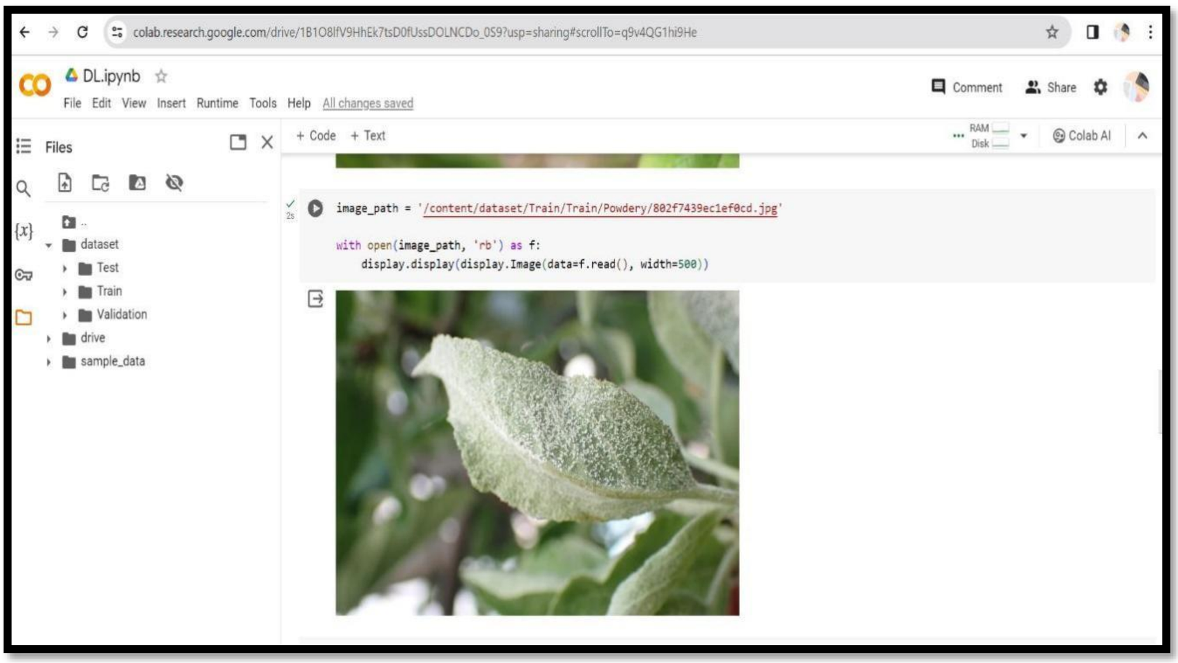

Fig6.4: Image of Powdery leaf in the 'Powdery' category of the dataset.

Here Figure 6.4 shows 'Powdery' category in the training set, illustrates a plant leaf affected by powdery mildew. Characterized by white, powdery patches, the visual indicators align with typical symptoms of this fungal infection. This image is crucial for training machine learning models to distinguish leaves with powdery mildew symptoms, contributing to a diverse dataset for plant health classification. The successful display in the output ensures accurate image path validation for effective model training and evaluation.

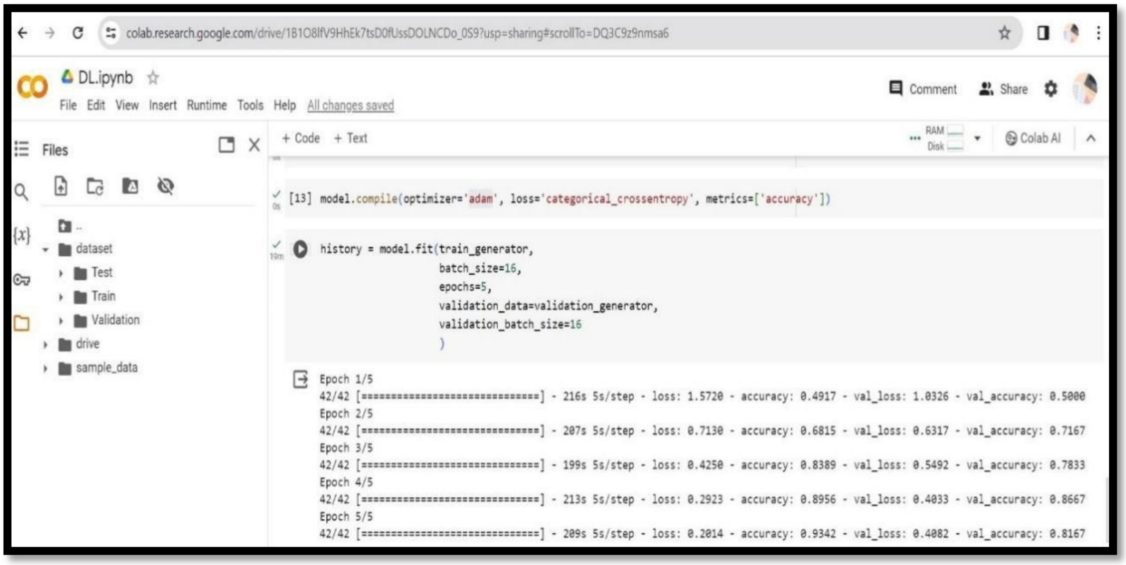

Fig 6.5: Model accuracy progress over five epochs, on both training and validation datasets.

The model fig 6.5 training history over five epochs reveals a progressive improvement in accuracy, starting at 49.17% and reaching 93.42% on the training data. Validation accuracy shows a similar upward trend, starting at 50% and peaking at 81.67%. The consistent reduction in loss values indicates effective learning.

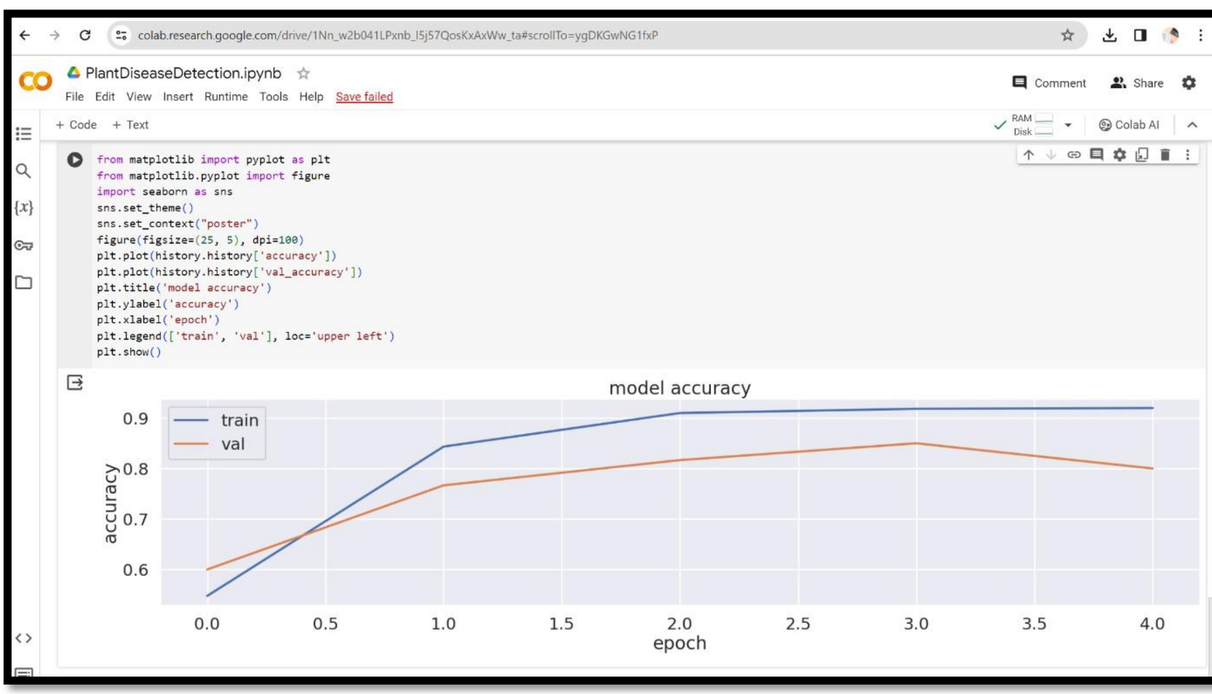

Fig 6.6: The graph shows the model's steady learning with a notable 93% accuracy on training and successful validation.

The plotted graph in figure 6.6 shows visually represents the training and validation accuracy of a model, showcasing the progression over epochs. The x-axis denotes the epochs, while the y-axis represents the accuracy percentage. The blue line represents the training accuracy, which steadily increases as the model learns from the training data, ultimately reaching an impressive accuracy of 93%. The orange line illustrates the validation accuracy, demonstrating how well the model generalizes to unseen data. The convergence of the two lines without signs of overfitting indicates a robust model. This graphical representation not only captures the model's learning trajectory but also highlights its ability to achieve a high level of accuracy on both the training and validation sets, signifying its effectiveness in the given task.

Conclusion

An extensive research study has been conducted, exploring a range of machine and deep learning techniques for the recognition and classification of plant diseases. Various classification methods within machine learning have been investigated to aid farmers in automatically detecting diseases across different crops. This analysis delves into diverse deep learning approaches for the detection of plant diseases, summarizing techniques and mappings for recognizing disease symptoms on plant leaves. The advancement of deep learning technologies in recent years has significantly contributed to the identification of plant leaf diseases, making this work a valuable resource for scientists engaged in plant disease detection.

FUTURE WORK

Future advancements in plant disease detection involve improving model accuracy and robustness, expanding detection capabilities to cover diverse plant species and diseases simultaneously. Real-time systems for prompt farmer intervention, integration with IoT and sensor networks, and user-friendly applications providing actionable insights are critical areas of development. Additionally, exploring transfer learning, enhancing model generalization, and collaborating with agricultural experts will contribute to the practical applicability and effectiveness of these detection methods in various agricultural contexts.

REFERENCES

[1] R. Jeevanantham, D. Vignesh, R. A. Abdul and A. J, "Deep Learning Based Plant Diseases Monitoring and Detection System," 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 360-365, doi: 10.1109/ICSCDS56580.2023.10104648.

[2] A Srivastava, A Pandey and S. K. Shukla, "Plant Disease Detection using SVM Algorithm and Neural Network Approach", International Journal of Research in Engineering and Science (IJRES) ISSN (Online): 2320-9364, ISSN (Print): 2320-9356.

[3] S. Garg, D. Dixit and S. S. Yadav, "Disease Detection in Plants Using KNN Algorithm," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 938-943, doi: 10.1109/ICAC3N56670.2022.10074491

[4] H. Cai and J. Jiang, "An Improved Plant Disease Detection Method Based on YOLOv5," 2023 15th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 2023, pp. 237-240, doi: 10.1109/IHMSC58761.2023.00062

[5] Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. Global Transitions Proceedings, 3(1), 305-310.

[6] Vasavi, P., Punitha, A., & Rao, T. V. N. (2022). Crop leaf disease detection and classification using machine learning and deep learning algorithms by visual symptoms: A review. International Journal of Electrical and Computer Engineering, 12(2), 2079.

[7] Sanida, T., Tsiktsiris, D., Sideris, A., & Dasygenis, M. (2022). A heterogeneous implementation for plant disease identification using deep learning. Multimedia Tools and Applications, 81(11), 15041-15059.

[8] Aggarwal, M., Khullar, V., & Goyal, N. (2022, October). Contemporary and Futuristic Intelligent Technologies for Rice Leaf Disease Detection. In 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-6). IEEE.

[9] Aggarwal, M., Khullar, V., & Goyal, N. (2022, October). Contemporary and Futuristic Intelligent Technologies for Rice Leaf Disease Detection. In 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-6). IEEE.

[10] Yu, Z., Zhang, L., Yuan, S., Bai, X., Wang, Y., Chen, X., ... & Zhou, X. (2023). Color ghost imaging through a dynamic scattering medium based on deep learning. *Optical Engineering*, *62*(2), 021005-021005.

[11] Pavithra, A., Kalpana, G., & Vigneswaran, T. (2023). Deep learning-based automated disease detection and classification model for precision agriculture. Soft Computing, 1-12.

[12] Joseph, D. S., Pawar, P. M., & Pramanik, R. (2023). Intelligent plant disease diagnosis using convolutional neural network: a review. Multimedia Tools and Applications, 82(14), 21415-21481.

[13] Shahi, T. B., Xu, C. Y., Neupane, A., & Guo, W. (2023). Recent Advances in Crop Disease Detection Using UAV and Deep Learning Techniques. Remote Sensing, 15(9), 2450.

[14] Dahiya, S., Gulati, T., & Gupta, D. (2022). Performance analysis of deep learning architectures for plant leaves disease detection. Measurement: Sensors, 24, 100581.

[15] Patil, R. R., Kumar, S., & Rani, R. (2022). Comparison of Artificial Intelligence Algorithms in Plant Disease Prediction. Revue d'Intelligence Artificielle, 36(2).

[16] Albattah, W., Nawaz, M., Javed, A., Masood, M., & Albahli, S. (2022). A novel deep learning method for detection and classification of plant diseases. Complex & Intelligent Systems, 1-18.

[17] Kurmi, Y., Saxena, P., Kirar, B. S., Gangwar, S., Chaurasia, V., & Goel, A. (2022). Deep CNN model for crops' diseases detection using leaf images. Multidimensional Systems and Signal Processing, 33(3), 981-1000.

[18] Daniya, T., & Vigneshwari, S. (2022). Deep neural network for disease detection in rice plant using the texture and deep features. The Computer Journal, 65(7), 1812-1825.

[19] Ashwinkumar, S., Rajagopal, S., Manimaran, V., & Jegajothi, B. (2022). Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks. Materials Today: Proceedings, 51, 480-487.

[20] Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. Global Transitions Proceedings, 3(1), 305-310.

[21] Guo, Y., Zhang, J., Yin, C., Hu, X., Zou, Y., Xue, Z., & Wang, W. (2020). Plant disease identification based on deep learning algorithm in smart farming. Discrete Dynamics in Nature and Society, 2020, 1-11.

[22] Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—a review. IEEE Access, 9, 56683-56698.