Leaf Disease Detection using Deep Learning

G. Arunalatha

Assistant Professor, Department of Computer Science and Engineering,
Perunthalaivar Kamarajar Institute of Engineering and Technology (PKIET),
Karaikal, Puducherry.
Email: vigneshgayu1121@gmail.com

Abstract

All plant species, whether wild or cultivated, can be affected by diseases. While each species has its own specific diseases to which it is vulnerable, the number of these diseases is generally limited. The incidence and frequency of plant diseases fluctuate with the seasons, influenced by factors such as the presence of pathogens, environmental conditions, and the types of crops and varieties being cultivated. Certain plant varieties are more prone to disease outbreaks, whereas others exhibit greater resistance. For thousands of years, humans have meticulously selected and cultivated plants for various purposes, including food, medicine, clothing, shelter, fiber, and aesthetics. Deep learning techniques have been effective in object detection and image classification. In this work, deep learning technology is utilized to detect the disease in the leaves of the plants.

Keywords: *Leaf, Deep learning, CNN.*

I. Introduction

Disease represents merely one of numerous risks that must be taken into account when plants are removed from their natural habitats and cultivated in pure stands under conditions that are frequently abnormal. Many valuable crops and ornamental plants exhibit a high susceptibility to disease and would struggle to survive in the wild without human assistance. Cultivated plants tend to be more vulnerable to diseases compared to their wild counterparts. This increased susceptibility arises from the cultivation of large populations of the same species or variety, which share a uniform genetic background and are grown in close proximity, sometimes spanning thousands of square kilometers. Under such circumstances, a pathogen can disseminate swiftly.

In general, a plant succumbs to disease when it is persistently affected by a causal agent that triggers an abnormal physiological process, disrupting the plant's normal structure, growth, function, or other activities. This disruption of one or more of a plant's vital physiological or biochemical systems leads to distinct pathological conditions or symptoms. Plant diseases can be broadly categorized based on the nature of their primary causal agent, which can be either infectious or noninfectious. Infectious plant diseases are instigated by pathogenic organisms such as fungi, bacteria, mycoplasmas, viruses, viroids, nematodes, or parasitic flowering plants. An infectious agent has the ability to reproduce within or on its host and can spread from one susceptible host to another. Conversely, noninfectious plant diseases arise from unfavorable growing conditions, which may include extreme temperatures, adverse moisture-oxygen relationships, toxic substances present in the soil or atmosphere, and an excess or deficiency of essential minerals. Since noninfectious causal agents are not organisms capable of reproducing within a host, they are not transmissible.

In nature, plants can be impacted by multiple disease-causing agents simultaneously. A plant facing a nutrient deficiency or an imbalance in soil moisture and oxygen levels is often more vulnerable to pathogen infections, and a plant already infected by one pathogen is frequently at risk of being invaded by secondary pathogens. The totality of all disease-causing agents affecting a plant constitutes the disease complex. To recognize a disease, it is essential to understand the normal growth patterns, varietal traits, and typical variability of plants within a species, particularly in relation to the environmental conditions in which they are growing.

The field dedicated to the study of plant diseases is known as plant pathology. The term pathology originates from two Greek words: pathos (meaning suffering or disease) and logos (meaning discourse or study). Therefore, plant pathology refers to the study of plant diseases. Fungi represent the most prevalent parasites responsible for plant diseases. Most fungi are microscopic organisms that can only be observed with a microscope; they feed on living green plants or decaying organic matter. When they infect living plants, diseases ensue. Fungi typically generate spores that can initiate an infection when they come into contact with a plant. These spores can be disseminated from one plant to another by means of wind, water, insects, and equipment. For fungus spores to initiate new infections, sufficient moisture and appropriate air temperature are necessary. Additionally, a wound on the plant may be required as an entry point for the fungus. Fungal diseases are particularly common during wet and humid seasons. Bacteria, which are single-celled microscopic organisms, can also attack living plants and induce plant diseases. Similar to fungi, bacteria can be transmitted from plant to plant through wind, rain splash, insects, and machinery. These diseases primarily manifest on leaves, although they can also affect stems and/or fruit. Leaf diseases are the most prevalent type of disease affecting most plants and are typically managed using fungicides, bactericides and resistant varieties.

II. Types of Leaf Diseases:

Leaf diseases are infections or disorders that affect plant leaves, often causing spots, discoloration, or deformities. They can be caused by fungi, bacteria, viruses, or even non-biological factors like nutrient deficiencies. Common symptoms include leaf spots, yellowing, wilting, and mildew.

1. Leaf Spots

Leaf spots, also known as anthracnose, scab, leaf blotch, or shot hole, typically present as distinct spots that vary in size, shape, and color. These spots usually have a clear margin. Occasionally, the spot, which may result from bacterial or fungal infection, is encircled by a yellow halo. When caused by a fungus, there is often some form of fungal growth within the spot, especially in humid conditions. This fungal growth can manifest as small, pimple-like structures, frequently black, or as a moldy spore growth. A hand lens or microscope is often required to observe these structures. If the spots are abundant or closely spaced, they may merge to create irregular areas referred to as "blotches." The common names for leaf spot diseases can be broad, such as bacterial leaf spot; descriptive, like frog-eye leaf spot; or named after the fungus, such as Septoria leaf spot.

Fig. 1 Leaf spot

Symptoms initially appear as small, dark purple to black spots on the leaf blade. As these spots grow, their centers often change to a light tan color. In warmer temperatures (above 85° F), the entire leaf blade may take on a dry, straw-like appearance. The disease primarily affects the leaf blades during cooler weather but can also infect leaf sheaths, crowns, and roots in hot, humid conditions.

Leaf spot is a disease that thrives in warm weather, yet the pathogen survives the winter as dormant mycelium in infected plants and decaying grass debris. The disease can become noticeable when temperatures reach 70° F (21.1° C). It is most severe when temperatures exceed 90° F and humidity levels are high. Drought stress followed by rewetting can exacerbate the disease.This disease is commonly found in home lawns and golf courses. General symptoms include spotted or wilted grass, while foliar symptoms consist of browning and spotting. These symptoms may manifest during the months of May, June, July, August, and September. The disease affects various hosts, including Kentucky Bluegrass, Perennial Ryegrass, Tall Fescue, Fine Fescue, Creeping Bentgrass, Annual Bluegrass, and Colonial Bent.

2. Leaf Blights

Leaf blights typically manifest as larger diseased areas compared to leaf spots and exhibit more irregular shapes. The blighted appearance of leaves may sometimes result from the merging of several small spots. Commonly, the term "blight" is included in the name, such as in Southern corn leaf blight or early blight.

Fig. 2 Leaf blight

Leaf blight denotes a widespread set of symptoms seen in various plant species and includes multiple crop diseases caused by different pathogens. Among the most prevalent symptoms of leaf blight diseases are leaf browning and desiccation, which can lead to diminished photosynthetic capacity if left untreated. Foliar blight infections have the potential to escalate into an epidemic, causing significant harm to photosynthetic functions and decreasing overall plant productivity. This can result in lower seed quality, impede plant growth and development, increase susceptibility to root rot, lead to stem death, and ultimately diminish yield. The economic repercussions vary based on disease prevalence, environmental factors, and the resistance of the specific plant cultivar. Leaf blight fungi and bacteria thrive in relative humidity levels exceeding 80% and temperatures ranging from 82 to 86°F (28 to 30°C). Fungal spores and bacteria can easily disseminate, carried by wind, water, insects, and agricultural tools to the soil. Once these pathogenic microorganisms encounter moisture, they proliferate rapidly within plant populations.

Leaf blight can inflict considerable damage on commercial farms, potentially resulting in yield losses of up to 50% in severe instances. Various fungi and bacteria are responsible for this condition, necessitating that growers accurately identify the specific pathogen to manage it effectively. Implementing sound cultural practices, such as utilizing pathogen-free seeds, practicing crop rotation, and ensuring optimal soil moisture, can significantly mitigate leaf blight. Additionally, applying fungicides at appropriate times is crucial in preventing the disease's spread. Digital farm monitoring platforms can facilitate the tracking of field conditions and the management of leaf blight treatments, allowing you to maintain control over the disease at all times.

3. Rusts

Rusts frequently create spots that resemble leaf spots; however, these spots are referred to as "pustules." Rust pustules can be bright yellow, orange-red, reddish-brown, or black. Typically, these pustules are elevated above the leaf surface, and when a white cloth is rubbed against them, a colored residue matching the color of the pustule is often visible on the cloth. In extreme cases, the affected leaf may wither and die quickly. Certain types of rust can also manifest on stems. Rusts are prevalent on grains and grasses.

Fig. 3 Rust

Rust diseases are prevalent fungal diseases in plants, classified under the order Pucciniales. They significantly affect the economy of several major crops. It is estimated that these diseases result in global losses of approximately $1 billion for wheat and $3 billion for coffee. Besides leaf rust, other forms of plant rust include stem and stripe rusts. Rust diseases are caused by various pathogenic fungi. For instance, Hemileia vastatrix is responsible for coffee leaf rust, Puccinia triticina causes wheat leaf rust, Puccinia striiformis leads to wheat stripe rust, and Phragmipedium spp. is responsible for common rust. Historically, rusts posed a significant threat to agriculture; however, advancements in plant science have improved disease management.

4. Powdery Mildew

Powdery mildew presents as a superficial, white to light grayish, powdery or mealy growth on leaves, but it can also affect stems and flowers. Leaves that are impacted typically turn yellow, wither, and die quickly. This issue is commonly found on cucurbit-type vegetables and small grains. Powdery mildew is a global disease that inflicts severe damage on crops and is regarded as more serious than downy mildew. The disease manifests in epiphytotic forms and is prevalent in temperate regions, particularly during dry weather, leading to significant crop losses. Approximately 10,000 species of angiosperms from over 1,600 genera are vulnerable to mildew fungi.

Fig. 4 Powdery mildew

Multiple genera can lead to the disease in various hosts. This condition is attributed to a group of fungi that belong to the order Erysiphales. The most prevalent genera include Erysiphe, Leveillua, Podosphaera, Sphaerotheca, and Uncincula. These represent the sexual stages, while in the asexual phase, the genus Oidium serves as the causal organism. The fungus acts as an obligate parasite, characterized by septate, hyaline, and extensively branched superficial mycelium that extends finger-like haustoria into the host cells. Powdery mildew creates a network of fungal filaments (hyphae) that are firmly anchored by specialized structures known as haustoria, which extract nutrients from the host. These hyphae can reproduce either sexually or asexually, resulting in the formation of cleistothecium or conidia, respectively.

5. Downy Mildew

Symptoms of downy mildew manifest as pale yellow-green to yellow patches on the upper surface of leaves, accompanied by light gray to purplish mold growth on the underside. The blue mold affecting tobacco is classified as a downy mildew disease. Additionally, downy mildew can lead to abnormal plant growth, often referred to as "crazy top," as observed in the case of sorghum downy mildew affecting corn or grain sorghum.

Fig. 5 Downy mildew

III. Literature Survey

Image processing and machine learning models [1] can be employed for the detection of plant diseases. Machine learning is a sub part of artificial intelligence which works automatically or give instructions to do a particular task. The main aim of machine learning is to understand the training data and fit that training data into models that should be useful to the people. The color of leaves, amount of damage to leaves, area of the leaf, texture parameters are used for classification. Random forest classifier has been used for classification.

A comparative analysis between support vector machines (SVM) model, K-Nearest Neighbor (KNN) model and convolution neural network (CNN) model is presented [2]. A deep convolutional neural network (CNN) is developed [3] based on a recently developed EfficientNet CNN model. The model was fine-tuned and trained for the detection of healthy and different unhealthy tomato leaf images.

A rice leaf disease detection system [4] using machine learning approaches is discussed. Three of the most common rice plant diseases namely leaf smut, bacterial leaf blight and brown spot diseases are detected. The dataset was trained on with a range of different machine learning algorithms including that of KNN(K-Nearest Neighbour), J48(Decision Tree), Naive Bayes and Logistic Regression. Decision tree algorithm, after 10-fold cross validation, achieved an accuracy of over 97% when applied on the test dataset.

In [5], leaf disease identification and grading using digital image processing and machine vision technology is devloped. . The proposed system is divided into two phases, in first phase the plant is recognized on the basis of the features of leaf, it includes pre-processing of leaf images, and feature extraction followed by Artificial Neural Network based training and classification for recognition of leaf. In second phase the disease present in the leaf is classified, this process includes K-Means based segmentation of defected area, feature extraction of defected portion and the ANN based classification of disease.

A system is designed for tomato leaf disease detection using the simplest approach while making use of minimal computing resources to achieve results comparable to state of the art techniques. Neural network models employ automatic feature extraction to aid in the classification of the input image into respective disease classes. This proposed system has achieved an average accuracy of 94-95 %.

IV. Deep learning:

Deep learning represents a branch of machine learning that employs multilayered neural networks, known as deep neural networks, to emulate the intricate decision-making capabilities of the human brain. Various forms of deep learning underpin the majority of artificial intelligence (AI) applications that we encounter in our daily lives. Deep neural networks are composed of several layers of interconnected nodes, with each layer enhancing the previous one to improve and fine-tune predictions or classifications. This sequence of calculations traversing the network is referred to as forward propagation. The layers at both the input and output of a deep neural network are termed visible layers. The input layer serves as the point where the deep learning model receives data for processing, while the output layer is responsible for generating the final prediction or classification.

a) CoAtNet:

Convolutional Neural Networks have emerged as a leading model architecture for computer vision since the advent of AlexNet. Following the success of self-attention models such as Transformers in the realm of natural language processing, numerous researchers have sought to harness the power of attention within computer vision. Recently, the Vision Transformer (ViT) has demonstrated commendable performance on ImageNet-1K. When pre-trained on the extensive JFT-300M dataset, ViT achieves results comparable to those of ConvNets, suggesting that it possesses equivalent capacity to convolutional networks. The integration of convolution and attention in machine learning is examined systematically from two fundamental perspectives—generalization and model capacity. The findings indicate that convolutional layers exhibit superior generalization, whereas attention mechanisms provide enhanced model capacity. By merging both convolutional and attention layers, we can attain improved generalization and capacity. CoAtNets is a hybrid model developed by Google's Brain Team that has recently attracted the interest of deep learning practitioners. It derives its name from the combination of two models: Convolution and Attention.

b) Architecture of AoAtNet:

For the convolution aspect, the MBconv blockis utilized, which is a type of image residual block designed with an inverted structure for enhanced efficiency, specifically for image models. It employs a narrow-wide-narrow strategy. For instance, we begin with a 1x1 convolution followed by a 3x3 depthwise convolution, resulting in a significant reduction in the number of parameters. In terms of the attention model, we implement the feed-forward network module (FFN module). The rationale for selecting the FFN module in the Transformer architecture, along with the "inverted bottleneck" of MBConv, is that it first expands the input channel size by a factor of 4, and subsequently projects the 4x-wide hidden state back to the original channel size to establish a residual connection. CoAtNet leverages the strengths of both Convolutional Neural Networks (CNNs) and Transformers, incorporating Translation Equivariance, Input-adaptive Weighting, and Global Receptive Field. These three characteristics are integrated seamlessly. Translation Equivariance: Convolutional neural networks possess a property known as Translation Equivariance, which indicates that the position of the image should not be fixed for the CNN to detect it effectively. Input-adaptive Weighting: This feature enables the attention model to discern the relationships among various elements in the input; however, it poses a risk of overfitting when the dataset

International Journal of Advance Research in Science and Engineering
Volume No. 14, Issue No. 06, June 2025
www.ijarse.com
IJARSE
ISSN 2319 - 8354

is limited. Global Receptive Field: In CNNs, the receptive field refers to the portion of the input matrix that impacts a specific unit within the network. Self-attention utilizes a broader receptive field, hence termed the Global Receptive Field, which allows the attention mechanism to acquire more contextual information compared to the conventional CNN receptive field.

The ideal basic block integrates the Input-Adaptive Weighting and Global Receptive Field characteristics of self-attention, alongside the Translation Equivariance found in CNNs. Once a manageable level of detail is achieved within a feature map, it is advisable to execute downsampling to diminish the spatial dimensions. Limit the global receptive fields G in attention to a local field L, akin to convolution. Instead of employing the quadratic Softmax attention variant, opt for a linear attention variant whose complexity is solely dependent on the spatial size. The down-sampling can be accomplished through two methods: 1) Convolution arises from expanded strides (for instance, strides of 16×16) as seen in ViT. 2) Multistage networks are characterized by gradual pooling, similar to ConvNets. By utilizing these options, we define a search space comprising 5 levels. The initial two layers, a traditional convolution and an MBConv, are utilized to decrease the image's dimensionality. The final three levels can incorporate either MBConv or the Transformer block. Consequently, there are 4 variants with progressively increasing numbers of Transformer stages: CCCC, CCCT, CCTT, and CTTT, where C and T represent Convolution and Transformer, respectively..

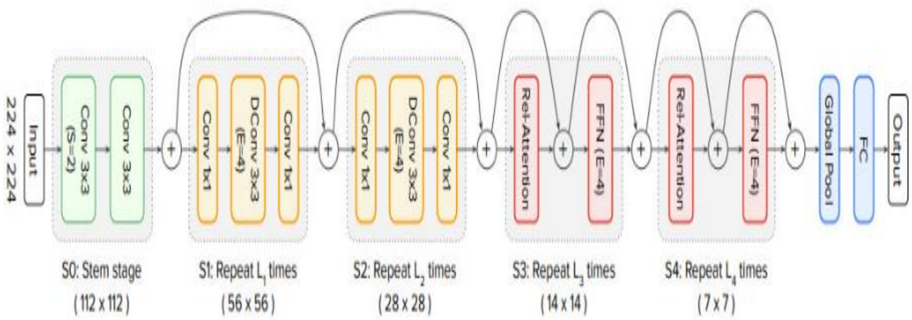

Fig. 6 CoAtNet Architecture

V. Dataset

a) Plant_leaves

This dataset consists of 4502 images of healthy and unhealthy plant leaves divided into 22 categories by species and state of health. The images are in high resolution JPG format.

b) Leaf

This dataset consists in a collection of shape and texture features extracted from digital images of leaf specimens originating from a total of 40 different plant species.

c) Kaggle Leaf Classification

The dataset consists approximately 1,584 images of leaf specimens (16 samples each of 99 species) which have been converted to binary black leaves against white backgrounds. Three sets of features are also provided per image: a shape contiguous descriptor, an interior texture histogram, and a fine-scale margin histogram. For each feature, a 64-attribute vector is given per leaf sample.

VI. Performance Metrics:

i. Accuracy

Accuracy is a fundamental evaluation metric for assessing the overall performance of a classification model. It is the ratio of the correctly predicted instances to the total instances in the dataset

Accuracy=TP+TN/TP+FP+TN+FN

ii. Precision

Precision evaluates the accuracy of the positive prediction made by the classifier. In precision specifies how many were actually positive in the total number of instances.

Precision = (TP) / (TP) + (FP)

iii. Recall

The recall is also known as sensitivity or true positive rate. It is the ratio of the number of true positive predictions to the total number of actual positive instances in the dataset. Recall measures the ability of a model to identify all relevant instances.

Recall = (TP) / (TP) + (FN)

iv. F1-Score

F1 score is the harmonic mean of precision and recall. It provide a single metric that balances the trade-off between precision and recall. It is especially useful when the class distribution is imbalanced.

F1 Score = 2 x [(Precision x Recall)/ (Precision + Recall)]

VII. Conclusion

Plant diseases pose a significant challenge to agricultural development in all nations. Timely and precise identification of leaf diseases is vital for mitigating economic impacts and enhancing crop productivity. Nevertheless, the reliance of farmers on traditional manual methods complicates the accurate identification of specific diseases. The early and precise detection and diagnosis of plant diseases are critical components of successful plant production. The manual identification of plant diseases is often a labor-intensive and error-prone endeavor. This approach can prove to be an unreliable means of recognizing and curbing the spread of plant diseases. The integration of advanced technologies such as Machine Learning (ML) and Deep Learning (DL) can address these issues by facilitating the early detection of plant diseases. In this study, a deep learning-based CoAtNet model is employed for the detection of leaf diseases.

References:

[1] Kulkarni, Pranesh, et al. "Plant disease detection using image processing and machine learning." arXiv preprint arXiv:2106.10698 (2021).

[2] Nikith, B. V., N. K. S. Keerthan, M. S. Praneeth, and T. Amrita. "Leaf disease detection and classification." Procedia Computer Science 218 (2023): 291-300.

[3] Chowdhury, Muhammad E. H., Tawsifur Rahman, Amith Khandakar, Mohamed Arselene Ayari, Aftab Ullah Khan, Muhammad Salman Khan, Nasser Al-Emadi, Mamun Bin Ibne Reaz,

Mohammad Tariqul Islam, and Sawal Hamid Md Ali. 2021. "Automatic and Reliable Leaf Disease Detection Using Deep Learning Techniques" *AgriEngineering* 3, no. 2: 294-312. https://doi.org/10.3390/agriengineering3020020.

[4] Ahmed, Kawcher, Tasmia Rahman Shahidi, Syed Md Irfanul Alam, and Sifat Momen. "Rice leaf disease detection using machine learning techniques." In 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1-5. IEEE, 2019.

[5] Rastogi, Aakanksha, Ritika Arora, and Shanu Sharma. "Leaf disease detection and grading using computer vision technology & fuzzy logic." 2015 2nd international conference on signal processing and integrated networks (SPIN). IEEE, 2015.

[6] P. Tm, A. Pranathi, K. SaiAshritha, N. B. Chittaragi and S. G. Koolagudi, "Tomato Leaf Disease Detection Using Convolutional Neural Networks," *2018 Eleventh International Conference on Contemporary Computing (IC3)*, Noida, India, 2018, pp. 1-5, doi: 10.1109/IC3.2018.8530532.