Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 1 (2024)

Plant Leaf Diseases Detection Using Image Processing Technique: A Review

[1]Ashutosh Shukla, [1]Rajan Kakkar

[1]*Lovely Professional University, Computer Science and Engineering*

Abstract: - Plant diseases caused by viruses pose a significant threat to agricultural productivity, affecting a wide range of plant species. With the increasing incidence of these diseases, the demand for rapid and accurate detection methods has grown. Image processing techniques have emerged as a potent tool for spotting disease symptoms in plants in response to this demand. This review paper focuses on the application of image processing techniques to precisely identify viruses in plant leaves, offering automated detection and texture value calculations. The paper presents accurate selection of research papers published over the past five years, chosen based on their titles, abstracts, and conclusions. These papers have been meticulously analyzed and included in this comprehensive review. The paper covers a number of topics related to plant disease detection, such as data sources, feature extraction strategies, pre-processing approaches, data enrichment strategies, and the use of various models for disease detection and classification. Additionally, it explores techniques for enhancing image quality and addressing overfitting issues, ultimately aiming to increase the precision of illness identification. Researchers can learn more about the possibilities of data-driven methods in the field of plant disease identification from this work, which is a great resource. By enhancing system performance and accuracy, the discussed techniques offer innovative solutions for addressing the impact of organisms on plant yields. Researchers and professionals in the agricultural domain will find this review paper a valuable reference in their pursuit of advanced methods for classifying and detecting plant diseases.

Keywords*: Plant disease identification, image processing technique, feature extraction, Deep learning, Machine learning.*

1. Introduction

Without the use of technology, farmers attempted to diagnose plant illnesses, which frequently resulted in a waste of time and money. Consequently, this had a detrimental impact on the country's economy. In response to this problem, we have turned to technology, specifically image processing, to identify and diagnose plant diseases. This technique involves capturing images of diseased plants and analyzing these images to identify the specific diseases affecting them. It provides a workable way to identify and diagnose diseases early on. Plant diseases can manifest in various ways, affecting different sections of the plant, such as the stems, roots, and leaves. Weather fluctuations can also contribute to the occurrence of diseases, further compromising crop health. This essay aims to explore the methods employed for the treatment of various plant diseases, with subsequent sections discussing different scenarios of disease identification. Plants can fall victim to diseases due to a combination of factors, including pathogenic microbes like fungi, bacteria, and parasites, soil pH, temperature fluctuations, moisture levels, humidity variations, and other environmental factors. These illnesses may significantly affect the growth of plants, vitality, and structural integrity, ultimately affecting the livelihoods of those who depend on these crops. Despite the challenges, many promising techniques are being developed to tackle this issue. One such advancement is deep learning, which has revolutionized agriculture by offering innovative approaches for plant disease identification. Using machine learning techniques, which are essential for automating the diagnosis of plant diseases, is another interesting approach. Moreover, efficient analysis and diagnosis of plant diseases is being achieved by the application of image processing techniques. These technologies provide a glimmer of hopefor maintaining crop health and food security by using photos of plant leaves to accurately identify and diagnoseplant illnesses [3].

A. Deep Learning Developments for Plant Disease Identification in Agriculture

Deep learning (DL), as a subset of machine learning (ML), has proven to be more responsive and effective. Unlike ML, where feature extraction and classification are separate steps, DL combines these tasks by employing multiple layers of processing. DL networks excel at recognizing complex patterns within the data, making them more suitable for handling unstructured data. Within the field of agriculture, DL methods including convolutional neural network (CNNs) and recurrent neural networks (RNNs) have gained prominence for identifying plant diseases from images of leaves. This paper reviews the latest research in this domain, including works by Turkoglu, who explored ResNet [36], Too, who delved into DenseNets [37], and Picon, who worked on ResNet-50 [38]. Other contributions, like Arsenovic's PlantDiseaseNet [39], X's DCNN [40], Pandian's Conv-5 DCNN [5], Jadhav's AlexNet [18], Sun's DM deep learning optimizer [24], and Akshai's DenseNet [22], all leveraged deep learning (DL) methods and convolutional neural networks (CNNs). One key advantage of CNNs is their minimal preprocessing requirements, as they excel in extracting features through the convolutional layer. This paper explores the pivotal role of deep learning in advancing agriculture's capacity, The figure1 below provides an overview of key aspects of DL in agriculture for plant disease identification."

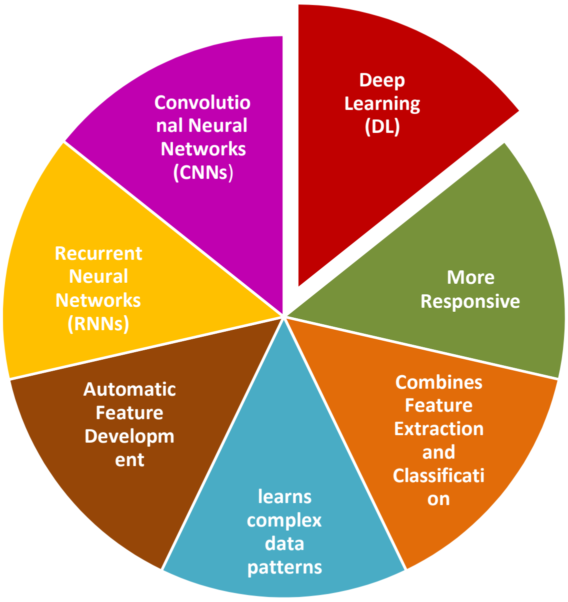

Figure 1: Overview of Deep Learning in Agriculture for Disease Identification.

B. Using Machine Learning Methods to Identify Plant Diseases

Machine learning (ML) has become a potent instrument, surpassing image processing, thanks to its use of loss functions. Loss functions enable ML models to make predictions about future behavior. When machine learning operates without explicit supervision, it excels in forecasting outcomes. ML-trained models become increasingly efficient over time as they accumulate experiences. The versatility of ML finds applications across diverse industries, from agriculture and from autonomous vehicles to digital assistants deep learning more responsive combines feature extraction and classification learns complex data patterns automatic feature development RNNs and CNNs such as Alexa and Siri, medical self-diagnosis, voice recognition, product recommendations, and more. Regarding the field of early diagnosis and categorization of plant diseases, ML methods open doors to innovative algorithms. This study delves into the most recent research in this domain and reveals the widespread adoption of ML techniques. For instance, Sharma's work on the CNN model [26] and Ghosh's investigation into CNN + kNN [9] underscore the utility of ML in plant disease detection. The accompanying figure 2 encapsulates the key features of ML demonstrating its function in the agriculture industry for the identification and categorization of diseases.

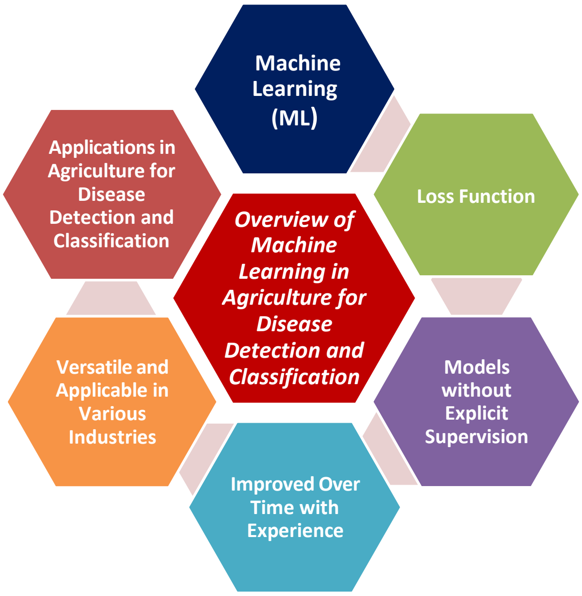

Figure 2: Overview of Machine Learning for Disease Classification and Detection in Agriculture.

C. Utilizing Image Processing Methods to Identify Plant Diseases

Image processing serves as a crucial tool for enhancing image quality and extracting essential data. This versatile technique finds applications in various fields, including medicine and agriculture, where it helps with tasks like pattern recognition, remote sensing, and color processing. The deployment of efficient and reliable image processing methods is instrumental in identifying diseases from plant leaf images. Image processing techniques encompass several key steps, including image capture, segmentation, classification, feature extraction, and image pre-processing. In this paper, we examine studies that use image processing methods to identify plant diseases. Notably, studies such as those by Malathy et al. [10] and Kumari et al. [41] have demonstrated impressive classification accuracy of up to 97% when diagnosing plant diseases, surpassing other approaches. The accompanying block figure 3 illustrates the multifaceted nature of overview of ML in Agriculture for disease detection and classification ML loss function models without explicit supervision improved over time with experience versatile and applicable in various industries applications in agriculture for disease detection and classification image processing in various industries, with a focus on disease detection in plant leaves and the critical steps involved in this process.

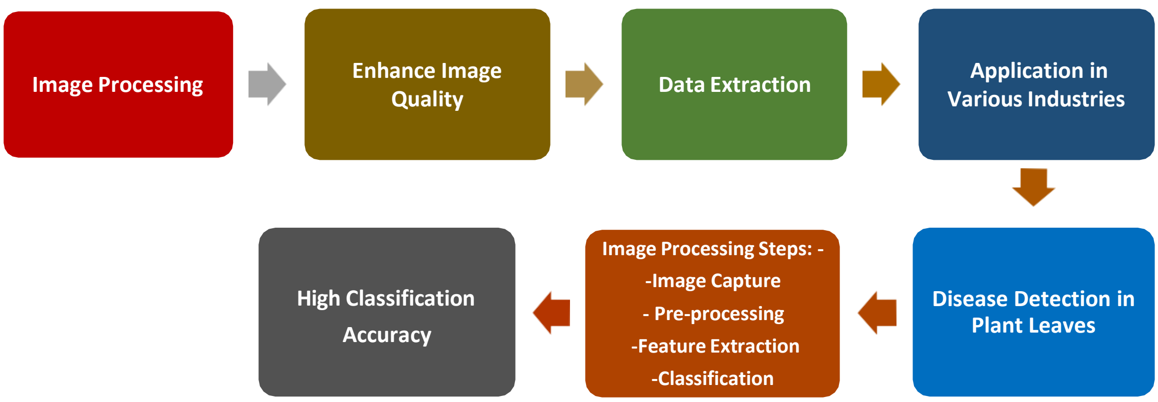

Figure 3: Overview of Image Processing Methods for Plant Leaf Disease Identification.

The introduction is the first of the six components that make up this essay, and the other sections are organized as follows. In the introduction section, you should provide the background information and context for your research and second is literature review the literature review, as discussed earlier, involves summarizing and analyzing existing research related to your topic. Emphasize important studies; the third portion is devoted to research methodology. Describe the strategies and tactics you employed to carry out your research in this area. Comparative Analysis is the fourth section. Within the section on comparative analysis, you typically compare and contrast different aspects of your research and section fifth is Result Analysis Here, present and analyze the results of your research. Use tables, graphs, and statistics to illustrate your findings and last section is conclusion and future scope The conclusion section should summarize the key findings of your study and their implications and in the future scope section, suggest directions for further research or extensions of your study.

2. Methodology

A methodology section in a review article on image processing-based leaf disease detection techniques would typically involve summarizing the different methods used. This section elucidates the strategies and criteria adopted for the selection of pertinent articles for this review.

2.1 Preparation Stage: -

The research process involved gathering articles from journals and conferences published between 2019 and 2023. The primary mode of collection was based on keyword searches in reputable databases like Google Scholar, SCOPUS Indexed Journal, and IEEE Xplore. From the aforementioned keywords, a total of [55] articles were initially retrieved. These underwent a screening process to narrow down relevant ones. It is display in table 1

Step	Description
Step1	Research period: 2019-2023
Step2	Data sources: IEEE Xplore, SCOPUS, Google Scholar
Step3	Keywords: [List of Keywords]
Step4	Initial Articles Retrieved: [55]
Step5	Screening Process
Step6	Relevant Articles: [45]

Table 1: Data Collection and Screening Process Overview.

2.2 Implementation Stage

The conducted phase involves a comprehensive review and synthesis of the selection standards for assessing current models, which include DL, image processing, and ML, particularly CNNs, in their effectiveness for detecting diseases in various crops and plants, utilizing diverse datasets. This flowchart is depicted through figure 4. A flowchart outlines the research methodology employed for this study as follows:

- **Start**
- **Find relevant papers:** Conduct keyword searches on IEEE Xplore, SCOPUS indexed Journals, and Google Scholar to identify pertinent research papers.
- **Filter for relevance:** Assess the identified papers based on their titles, abstracts, conclusions, and full texts to determine their relevance to the research needs.
- **Is the paper relevant?**
  - **Yes:** Proceed to the next steps.
  - **No:** End.

- **Identify techniques:** Examine the detection and classification techniques utilized in the selected papers.
- **Analyze outcomes:** Investigate the performance metrics and results achieved by these models.
- **Summarize findings:** Write a synopsis and consolidation of the most important conclusions and revelations from the literature review.
- **Organize reviewed literature:** Arrange and categorize the literature to create a structured review.
- **Stop**

An overview of the use of ML for disease classification and detection in agriculture. The full research process used in this study to assess disease detection algorithms in crop and plant datasets is represented by this flowchart.

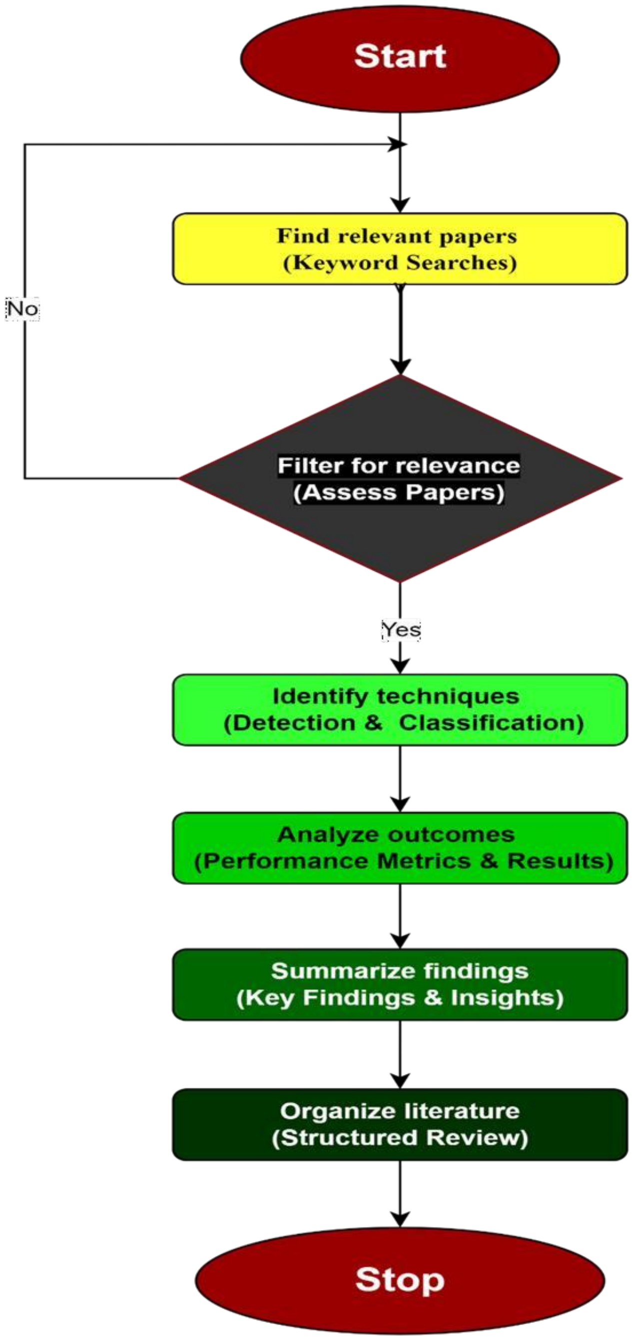

Figure 4: Crop and Plant Disease Detection Model Evaluation Flowchart.

2.3. Literature Search and Selection Process

An initial extraction using the above keywords yielded 45 articles focusing on plant disease detection and categorization. These five-year (2019–2023) publications were gathered from IEEE Xplore, SCOPUS Indexed Journal, and Google Scholar. The sifting process was tripartite. Firstly, based on the titles, the number was refined to 50 articles. Subsequent scrutiny of abstracts and conclusions further reduced the count. A thorough perusal of the full text of these papers finally left us with 42 pertinent articles. The distribution of these articles over the years (2019-2023) can be observed in figures 5 and 6. The goal was to craft a comprehensive systematic review based on these findings.

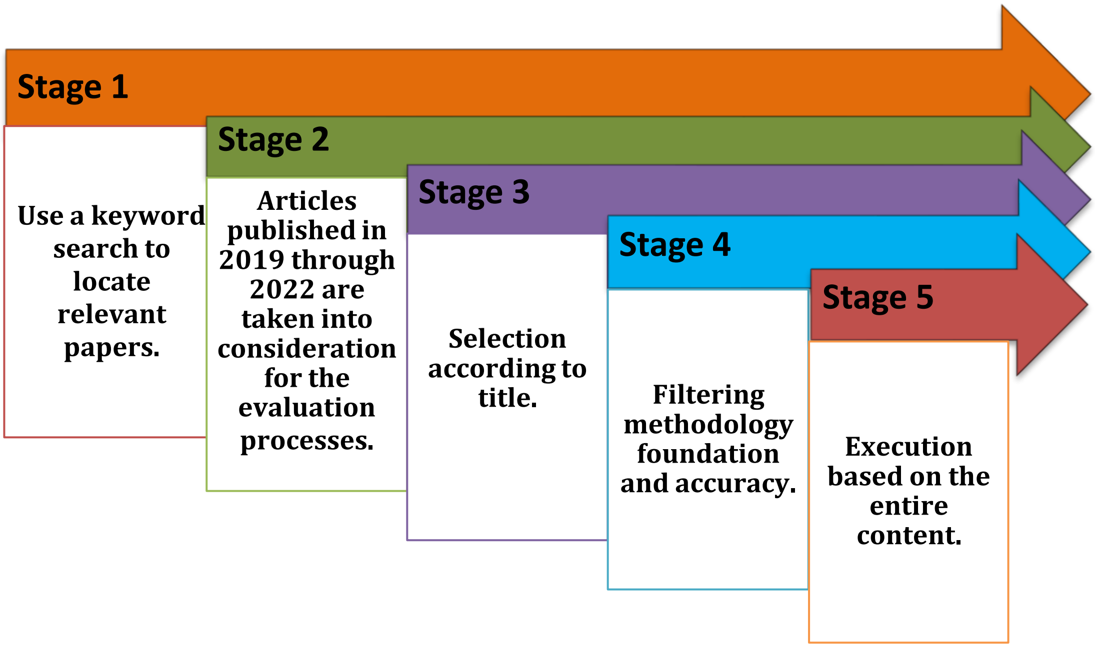

Figure 5: Shows the criteria for inclusion and exclusion that were applied to this study.

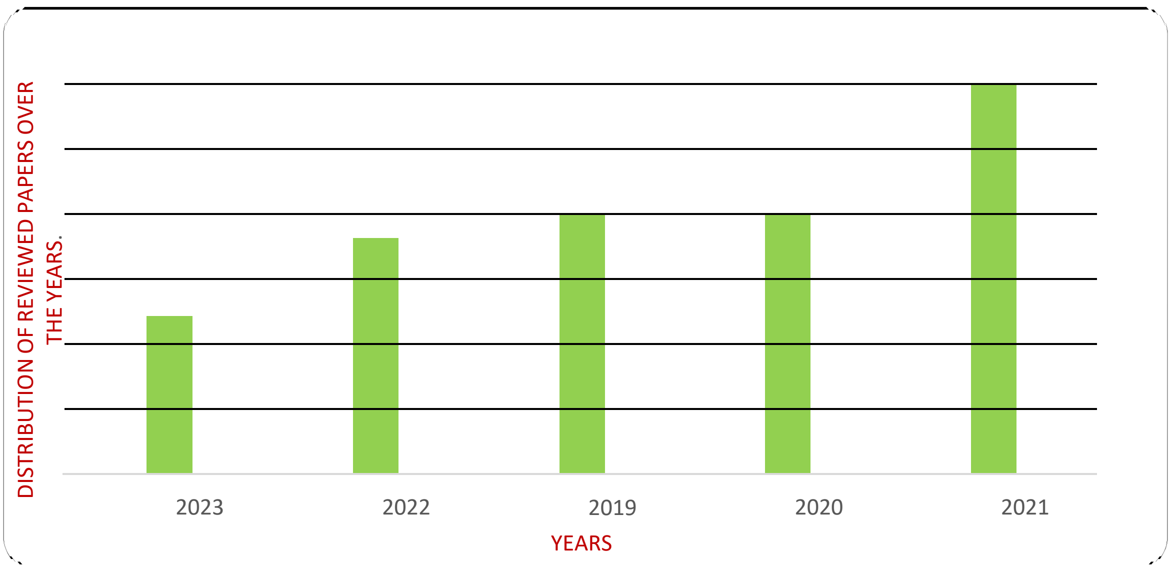

Figure 6: Displays the annual number of papers from 2019 to 2022.

3. Literature review

In this comprehensive literature review, we delve into various critical aspects of plant disease detection usingimage data. We begin by examining the diverse sources researchers have tapped into for collecting plant image data, shedding light on the breadth of data acquisition in this field. Next, we investigate the pre- processing methods used to improve the picture data's quality, ensuring it is optimal for further analysis. Dataaugmentation techniques are then discussed, providing insights into how researchers expand their datasets toimprove model performance. Feature extraction methods take center stage, where we investigate the approaches used to extract essential information from plant images. The review proceeds by concentrating on the models used to diagnose and classify plant diseases, emphasizing the critical functions of DL and ML.Image quality improvement techniques are explored, offering solutions for refining the visual data. Strategiesfor reducing overfitting in plant disease detection models are discussed, addressing a common challenge in this field. We then delve into the specific plant species and disease classes evaluated in the studies, showcasing the diversity of plant-related research. Finally, we assess the crucial aspect of accuracy in plant disease identification, summarizing the core findings and contributions of the literature. This comprehensivereview provides valuable insights into the multifaceted landscape of plant disease detection using image data.After a thorough analysis of multiple research papers, we have structured this literature review into distinct section as depicted in figure 7.

Figure 7: Comprehensive Framework for Plant Disease Detection Pipeline.

1. Data Acquisition Sources in Plant Image Data Collection

The initial stage in identifying and classifying leaf diseases involves obtaining data in the form of images of plant leaves. This section discusses the various sources from which researchers collect plant image data. It covers datasets such as Kaggle's Plant Village, real field environments, publicly available datasets, specific plant species data, hyperspectral UAV images, and other unique sources. These sources serve as the foundation for plant disease detection studies, influencing the quality and diversity of the data. The sources from which data has been obtained for the literature are visualized in Figure 8, This chart provides a breakdown of the various origins of data used in the studies."

In the study by S. Phani P, et al. [1], The "Plant Village dataset," which provides a collection of 4,062 photos of grape leaves displaying signs connected to various ailments such as Black Rot, Esca measles, Leaf spot, and healthy states, was used to detect grape leaf diseases. X., Zhang et al. [2] underscores the substantial impact of plant diseases on agricultural production and the critical need for early detection to mitigate this impact. The research predominantly relies on the 'Plant Village' dataset, sourced from Kaggle, with a particular focus on potato leaves. This dataset provided a comprehensive collection of images for studying various plant diseases. Mahum et al. [4] utilized the "The Plant Village" dataset in combination with manually collected data to gather relevant images of potato leaves, aiming to create a comprehensive dataset for disease detection. Pandian et al. [5] A set of about 55,448 photos showing both healthy and unhealthy leaves from different plant types were acquired. These pictures are from a dataset that was made available to the public. Vallabhajosyula et al. [6] We collected RGB photos of fourteen different crop varieties and found 38 different kinds of leaf damage. These images were sourced from the PlantVillage dataset, a readily available resource accessible through Kaggle. Ashwinkumar et al. [7] Their plant image data came from PlantVillage, an open-source dataset that was obtained via a download from Kaggle. This dataset included 38 classes of damaged leaves and RGB photos of 14 different crop species. A sizable dataset consisting of 1500 potato leaf photos and 5932 rice leaf images was gathered by Sharma et al. [8] from reliable agricultural databases. Their deep learning model was trained and tested using these sources as a basis. Ghosh, S et al. [9] three different datasets—LeafSnap, Flavia, and MalayaKew—to perform their research,highlighting the value of a variety of trustworthy sources when gathering plant image data. Chelleapandi et al.
[14] made use of the 54,306 photos in the Plant Village Dataset, which depicts 14 plant species and 38 disease types. The Git-Hub repository of SP Mohanty provided the dataset. jain et al. [17] They used the Kaggle-hosted new plant diseases dataset to source their image data. The chosen photos show three distinct crops, rice, grapes, and maize, illustrating the wide range of plants examined for disease detection. In Jadhav et al.'s [18] research, picture data connected to soybeans from several soybean farms in Kolhapur district, Maharashtra, India, served as the main source of data for gathering information about plants. This demonstrates how actual environmental data collection may be used to produce a dataset that nearly mimics the circumstances of soybean farms in the Kolhapur area. Chowdhury et al. [19] mainly used the PlantVillage dataset, which gave them access to about 18,100 photos of tomato leaves, for their research. The dataset included a thorough set of ten classes, nine of which represented different types of damaged tomato leaves and one representing healthy leaves. In their study, Lijo, J. et al. [20] gathered a sizable dataset of 10,000 photos showing both healthy and sick leaves. The photographs were taken from the Plant Village collection, which included a wide range of plant species, includingtomato, pepper, potato, mango, strawberry, and grape. Their research benefited greatly from this varied dataset, which included information on a variety of bacterial and fungal infections. Sun et al. [24] gathered their information from the publicly available PlantVillage dataset. For their investigation, this dataset was an invaluable source of plant photos. Eighty percent of the data were put aside for training and twenty percent were set aside for testing the suggested model. Sujatha et al. [16] collected their data from citrus leaves in a real environment ata citrus research center in Sargodha City, using DSLR cameras with a 72-dpi resolution, under expert guidance. This source of data is valuable for the study. In the study conducted by Abbas et al. [21], the primary source of data was the open-source PlantVillage dataset. This dataset provided them with both healthy and diseased tomatoleaf images, offering a comprehensive dataset to train and test their models. Akshai et al. [22] relied on the PlantVillage dataset as their primary source of data for their research. From this dataset, they acquired a total of 4060 grape leaf images, encompassing both healthy specimens and those afflicted with various diseases. Umit

Atila et al. [23] utilized the PlantVillage dataset for their research, a comprehensive resource containing 55,448 original images from 14 different plant species, both healthy and diseased. Wang et al. [25] utilized the PlantVillage dataset as their primary source of data. This dataset comprises 3000 plant leaf images and is distinguished by its inclusion of 15 subcategories, encompassing diverse plant species and various disease severity levels. Such comprehensive datasets are pivotal in enhancing the robustness of disease detection models. Sharma, P et al. [26] sourced their image data from various places, including the Plant Village database, the internet, and local farms. In total, they collected 17,929 images, with a specific focus on tomato plants and 10 disease classes. Karthik R et al. [27] utilized the Plant Village Dataset for their research. This dataset contains images representing three tomato leaf diseases. Dengshan Li et al. [28] collected a dataset for their research, focusing on identifying rice plant diseases and pests. This dataset was gathered from provinces in China between June and August 2018. It encompassed images captured using mobile phones and a specific Sony WIFI-controlled camera, as well as videos recorded with an iPhone7, with the primary goal of addressing crop-related issues. The source of plant image data used in this research by Nithish, E.K et al. [30] is not explicitly mentioned in the provided paragraph. The study aimed to enhance the dataset for detecting diseases in tomato leaves. Chen, J et al. [31] obtained their dataset from Xiamen, China's Fujian Institute of Subtropical Botany. This large dataset (about 1000 photos) showed the many illnesses that affected the leaves of maize and rice. The photos, which were taken in actual environments, showed the leaves in various lighting scenarios over a background of crowded fields. In Pham, T.N, et.ai.'s research [33], Approximately 450 photos of mango leaves were captured with a 3096 × 3096 resolution camera in Giang Province, Vietnam. Images of both healthy and diseased mango leaves were includedin this dataset. Marzougui, F et.ai. [34] collected 500 photos in total, 250 of which showed healthy leaves and the remaining 250 of which showed damaged ones. The data was obtained by the researchers directly from genuine situations. The Samsung Intelligent LCD camera was used to take the pictures. Selvam, L et.ai. [35] gathered over 1085 photos of ladyfinger leaves, including those that were disease-free, ill from overuse of fertilizer, and healthy leaves. These photos were taken with an 8 MP Samsung A7 smartphone camera, straight from fields in the Tiruvannamalai district. M, Turkoglu et al. [36] collected information from actual field settings, concentrating on a variety of species such as cherry, walnut, apricot, and peach. In their thorough investigation, Too, E.C et al. [37] categorized and identified plant illnesses in a variety of plant species using the well-known PlantVillage dataset. In the research conducted by Picon, A. et al. [38], To comprehend and recognize certain plant diseases, a real-world dataset focused on wheat plants was used. The study conducted by Arsenovic, M. et al. [39] used real-world datasets that included photos of different plants taken in an agricultural field setting. The study conducted by X., Zhang et al. [40] using hyperspectral UAV photos and real-world datasets to detect yellow rust in wheat plants. The study conducted by Al Haque et al. [42] gathered a thorough dataset of ten thousand guava photos taken with a Nikon D7200 DSLR camera in a variety of settings. These photos showed four different conditions: healthy guava, fruit rot, anthracnose, and fruit canker. The information came from actual settings as well as, to some degree, the internet. M. Francis et al. [43], Pre-processing involved cropping and downsizing the photos to 64 x 64- pixel proportions. These methods set up the pictures for examination. In this study by Sahithya, V. et al. [44], Images of lady fingers were taken in a real-world setting with a digital camera with a dimension of 1584 x 3456.

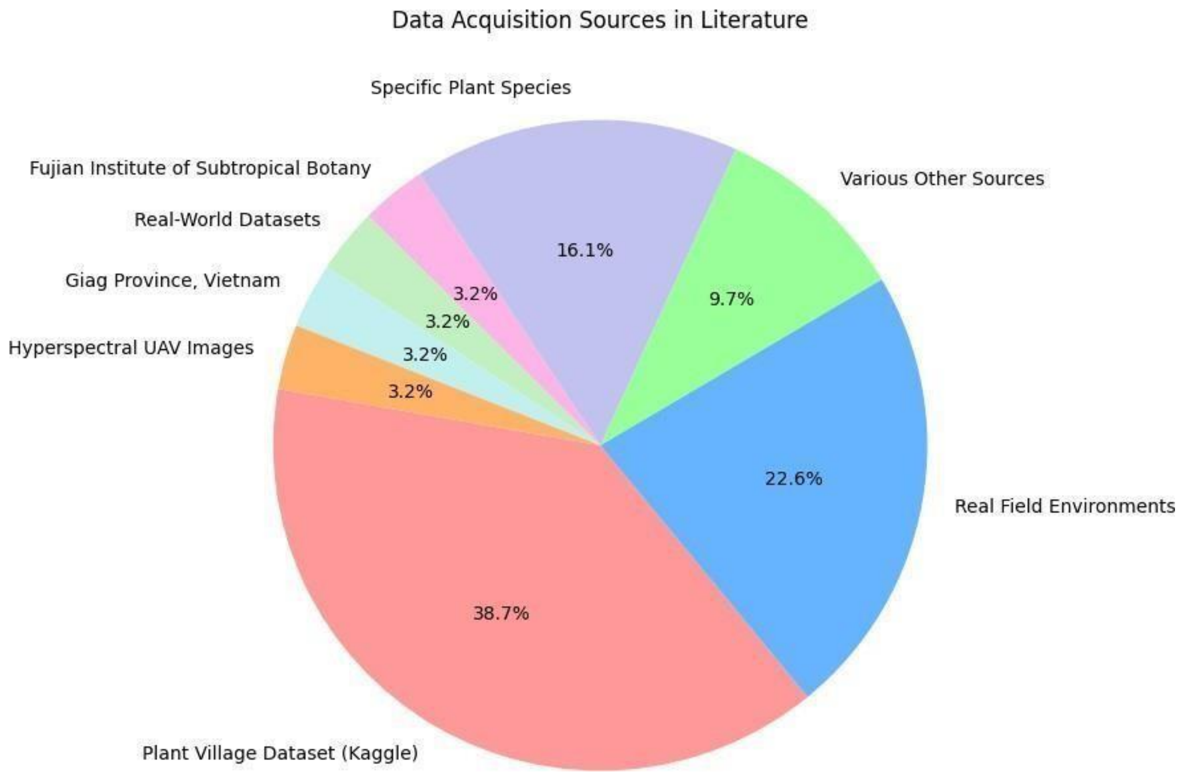

Figure 8: Distribution of Data Acquisition Sources.

2. Pre-Processing Techniques for Plant Image Data

To prepare the image data for subsequent processing, several pre-processing techniques were applied. Technologies for pre-processing are crucial for getting raw picture data ready for analysis. Researchers employ methods like non-linear filtering, image resizing, enhancement, Gaussian filtering, image shearing, zooming, flipping, and more to enhance image quality, reduce noise, and ensure data cleanliness. Figure 9, illustrates the utilization of various pre-processing techniques in the analyzed studies. Each horizontal bar represents a specific pre-processing method employed in the research. The bars are color-coded in a calming sky blue. The length of each bar corresponds to the number of studies that incorporated the respective technique. Notably, the y-axis is inverted, placing the most frequently used technique at the top for easy reference. This visual representation provides a clear overview of the prevalence of different pre-processing techniques within the analyzed studies, aiding in the understanding of the techniques' adoption in this research context."

When preprocessing is underway, Ashwinkumar et al. [7] employed non-linear filtering techniques, specifically bilateral filters, to enhance image quality by reducing noise in the acquired images. The image data was pre-processed by Malathy et al. [10] using methods for restoring the image and scaling. These processes aim to reduce image noise and enhance the sharpness, ensuring better quality images are available for the subsequent stages of their research. Wassan et al. [11] gathered historical weather data for their study. The data, which covered the months of January 2015 through February 2019, was gathered from several US weather stations. While this data is primarily weather-related, it is crucial for understanding the impact of frost events on plants. Chouhan et al.'s study [13], pre-processing of image data was a key focus. They improved the quality of the image data by using a variety of procedures like scaling, restoration, and image enhancement, readying it for further study. In order to guarantee that the data is clean and prepared for feature extraction and modeling, certain pre-processing procedures are essential. Chelleapandi et al. [14] involved rigorous pre-processing procedures, including image shearing, zooming, flipping, and brightness adjustment, conducted via the Image Data Generator. These techniques enhanced the quality and diversity of the image dataset, preparing it for deep learning analysis. In their research, Jain et al. [17] employed a 3 × 3 Gaussian filter during the pre-processing phase. This method effectively eliminates noise from the images, ensuring that subsequent analysis is based on clear and distinct image features. The pre-processing techniques Jadhav et al.'s [18] required resizing the pictures to fit into two distinct, specially designed proportions for GoogleNet and AlexNet. In particular, 550 photos of soybean leaf samples were processed to dimensions of 224 × 224 × 3 for GoogleNet, and 649 photographs of soybeans were downsized to 227 × 227 × 3 for AlexNet. These dimensions highlight how crucial it is to prepare images specifically for a given deep learning model. Chowdhury et al. [19] Pre-processing methods used in this investigation included normalization and scaling. For EfficientNet and U-net segmentation algorithms, the images were enlarged to 224 × 224 and 256 × 256 in dimensions, respectively. To improve data consistency and quality, z-score statistics were also used to standardize the picture data. Lijo, J. et al. [20] The obtained photos were scaled to a standard dimension of 256 × 256 pixels during the pre-processing phase. In order to guarantee consistency and compatibility in further processing and analysis, this step is essential. Chen, J et al. [31] To streamline the dataset,a pre-processing step was undertaken to standardize the shorter sides of the images, ensuring uniformity in imagesize. Pham, T.N, et.ai. [33] The obtained images of mango leaves were downscaled to a more reasonable resolution during the preprocessing stage. Furthermore, pixel intensities were adjusted using contrast enhancement techniques, which raised the overall quality of the images for further study. Kumari, C.U. et al. [41] involved converting plant images to grayscale and segmenting clusters, particularly focusing on the leaf segments affectedby diseases. Al Haque et al. [42] Pre-processing techniques were employed to prepare the collected guava imagesfor analysis. In this study by Sahithya, V. et al. [44] Pre-processing involved resizing the images to a consistent standard size. In the research by M. R. Howlader et al. [45], Python code was utilized in the pre-processing stageto resize every image collected to a uniform size of 256 by 256 pixels.

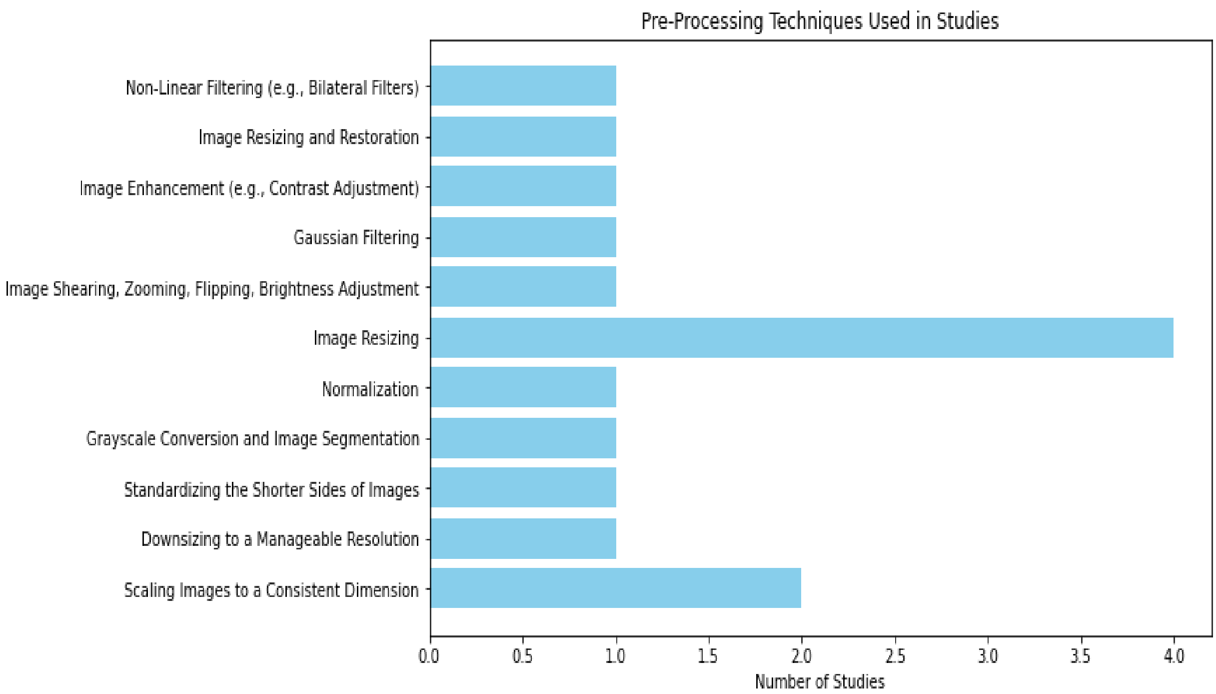

Figure 9: Pre-Processing Techniques Used in Plant Image Data.

3. Data Augmentation Techniques for Expanding Datasets

To overcome limited data challenges, researchers use data augmentation techniques to increase dataset size. Techniques include neural style transfer, adjustments to image position and color, GANs, PCA, scaling, rotation, flipping, SMOTE, and more, contributing to dataset diversity and mitigating class imbalances. This section explores different methods for increasing the number of images in a dataset, a practice known as data augmentation. Researchers have employed various techniques to expand their datasets, and this section provides an overview of these methods. To illustrate how various data augmentation strategies have been applied in the papers we've analyzed, we've produced a pie chart. Each Augmentation method is represented as a slice within the pie chart, and the size of each slice indicates how many studies have utilized that specific technique. The chart, complete with a helpful legend, is labeled as Figure 10, It provides a straightforward way to comprehend the distribution of data enhancement techniques within the studies we've examined. This visual representation simplifies the process of grasping how these techniques have been applied across the various studies.

In this research, by X., Zhang et al. [2] used a sophisticated convolutional neural network (CNN) model designed specifically for the effective detection of leaf diseases. They addressed the challenge of limited data by utilizing data augmentation techniques. The dataset that contained 1700 photos of potato leaves was used to train the model. Pandian et al. [5] Various image augmentation techniques were applied to enhance the dataset. These methods included principal component analysis, deep convolutional generative adversarial networks, neural style transfer, and color and position modifications for images. These methods were employed to expand the initial 55,448 images into a more extensive dataset containing 234,008 images. Vallabhajosyula et al. [6] to bolster the dataset's size and mitigate overfitting, we employed four distinct data augmentation methods. These comprised methods for picture augmentation, translation, rotation, and scaling. Chelleapandi et al. [14] The ImageDataGenerator program was essential to data augmentation since it allowed for a variety of modifications, including flipping, shearing, zooming, and brightness adjustment, which increased the dataset and guaranteed the stability of the model. Chowdhury et al. [19] Three affine transformation procedures were used as part of the data augmentation techniques: translation (both vertically and horizontally), rotation (both clockwise and counterclockwise), and scaling. These methods increased the size of the image dataset, which strengthened the model training. Lijo, J. et al. [20] used a number of data augmentation methods to increase the variety of the dataset and lower the overfitting risk. These methods included noise reduction, brightness enhancement, contrast enhancement, and rotation. Abbaset al. [21] underlined how important it is to enhance data in order to prevent overfitting. They added artificial yetrealistic tomato leaf images to their original dataset by using the Conditional Generative Adversarial Network (C-GAN) to create synthetic images. Akshai et al. [22] Recognizing the need for a more diverse dataset to enhance the robustness of their models and mitigate overfitting, Akshai et al. employed several data augmentation techniques. Specific strategies included image rotation, shifting, and zooming to artificially expand their dataset without compromising the integrity of the original images. Umit Atila et al. [23] augmented the original dataset, expanding it from 55,448 images to 61,486 images. This amplification likely involved techniques such as rotation,flipping, and other variations to diversify the dataset, which enhances model generalization. Sharma, P et al. [26]To address the homogeneity of the Plant Village dataset, the researchers introduced data augmentation techniques. These techniques included random transformations like brightness adjustment, contrast modification, and blur. The purpose of these augmentations was to enhance the dataset's diversity and improve the model's performance.Ji, M et al. [29] used a variety of data augmentation techniques, such as rotation, zooming, flipping, shearing, andcolor changes, to enhance their dataset. By using these strategies, the dataset's variety was increased, giving the model more diverse training samples. Nithish, E.K et al. [30] applied two primary data augmentation techniques:random resized Crop, which adjusted image sizes, and random rotation at 30 degrees. These techniques expanded the dataset fourfold, increasing the diversity of examples available for model training. Chen, J et al. [31]employed data augmentation techniques to increase the dataset's volume and variety. Implemented techniques included rotation, flipping, scaling, and translation. Beyond merely expanding the dataset, these augmentations also aimed to combat the prevalent challenge of model overfitting. Marzougui, F et.ai. [34] To bolster the datasetand enhance model performance, data augmentation techniques were employed. Marzougui and the team expanded the dataset using Kera's Image Data Generator class, applying operations such as horizontal flipping, 30-degree rotations, and both horizontal and vertical shifts. The "nearest" setting was selected as the fill mode to handle areas outside the input boundaries, effectively augmenting the data. To enrich the dataset and enhance the model's robustness, Selvan [35] employed five data augmentation techniques: rotation, horizontal flipping, shearing, zooming, and shifting in both height and width directions. These techniques increased the diversity and volume of ladyfinger leaf images for comprehensive analysis. Al Haque et al. [42] In order to increase the dataset and lessen the overfitting of the model, data augmentation techniques such as horizontal flipping, zooming, height and breadth shifting, rotation, nearest fill, and shearing were applied.

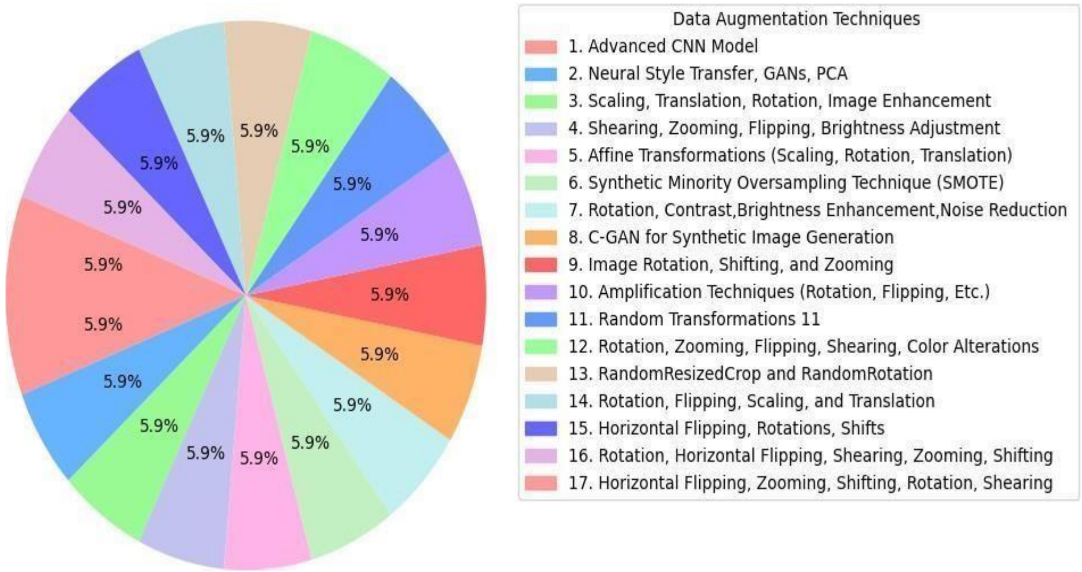

Figure 10: Distribution of Data Augmentation Techniques in Literature.

4. Feature Extraction Techniques in Plant Disease Detection

This section explores feature extraction methods used in plant disease detection studies. Techniques range from employing advanced CNN models like DenseNet and MobileNet to using Residual Networks, attention mechanisms, custom DCNN architectures, and statistical measures for characterizing image clusters. Feature extraction is crucial for disease classification to illustrate the distribution of different feature extraction methods used in the field of plant disease diagnosis; a pie chart is created. Each technique is represented as a segment within the pie chart, and the size of each segment is proportional to the number of studies or applications that have incorporated that specific technique. The chart is appropriately titled "Feature Extraction Techniques in Plant Disease Detection" to offer context, and it serves as an informative visual representation of the prevalence of these techniques in the field. This visualization helps viewers discern which techniques are more commonly employed in plant disease detection applications. The chart is presented as Figure 11, to further facilitate reference and understanding.

Mahum et al. [4] utilized the harness the power of an Efficient DenseNet model, which inherently focuses on feature extraction for disease classification. Pandian et al. [5] The feature extraction process made use of carefully selected convolutional layers optimized for the task. To extract essential data from plant images, Ashwinkumar et al. [7] employed the MobileNet model, a well-known CNN-based strategy. Ghosh, S et al. [9] uses CNNs to extract features, showing how deep learning methods can be applied to tasks involving the classification of plants. In Chen et al.'s [15] research, various feature extraction methods were explored, with an emphasis on feature extraction using RESNET18, a convolutional neural network. This emphasizes how important feature extraction methods are when it comes to identifying plant diseases. Wang et al. [25] introduction of the DBA_SSD model, which integrates a residual network and an attention mechanism, offers a refined approach to feature extraction. This novel architecture ensures that significant features from plant images are extracted, promoting accurate disease detection. Karthik R et al. [27] employed CNN for the automatic extraction of relevant features for disease detection. They presented two distinct deep architectures to achieve this. The first relies on residual learning coupled with a CNN, and the second enhances the first by integrating an attention mechanism. This attention mechanism aids the model in focusing on specific areas of the image, presumably where disease symptoms are more prominent. Dengshan Li et al. [28] made use of a unique deep convolutional neural network (DCNN) architecture to extract relevant characteristics from movies and photos. Kumari, C.U. et al. [41] used statistical metrics like energy, correlation, variance, mean, contrast, standard deviation, and homogeneity along with feature extraction techniques to describe the segmented clusters in plant photos. Sahithya, V. et al. [44] used the Grey Level Co-occurrence Matrix (GLCM) approach to extract features. This method included elements of texture, color, and geometry. Chouhan et al [13].'s research likely extracted features relevant to plant disease detection, which can include texture, shape, or color-based features. Jain et al. [17] mainly concentrated on removing two kinds of features—texture and color—from photos of plants. While color features were calculated utilizing factors such as skewness, standard deviation, kurtosis, and mean from the color moment equation, texture features were obtained using a Gray-Level Co-occurrence Matrix (GLCM). This emphasizes how crucial feature extraction is for identifying plant diseases.

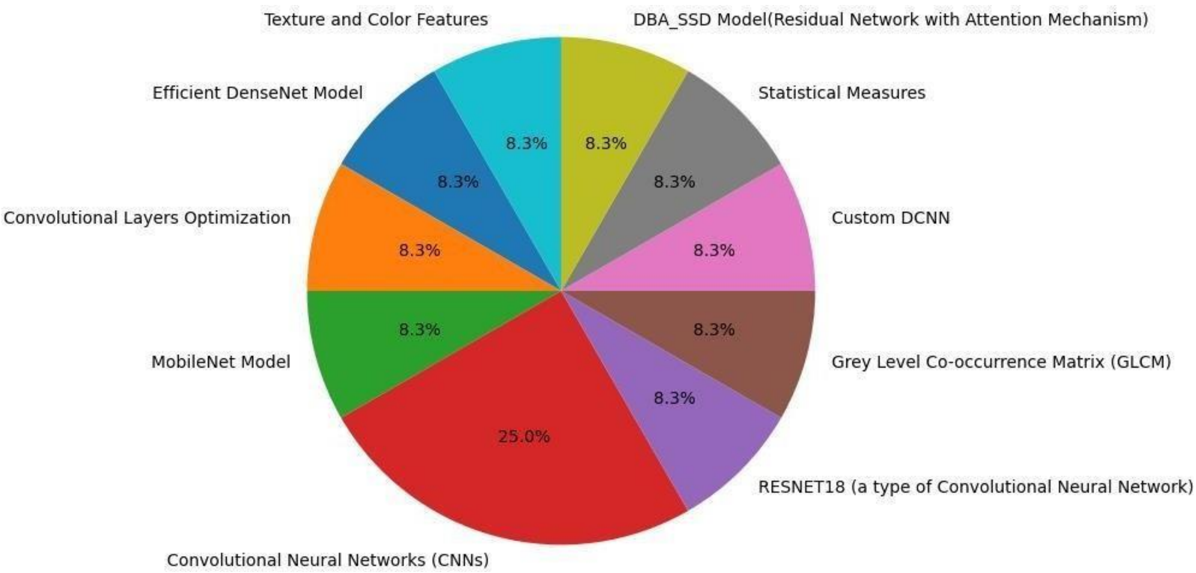

Figure 11: Dispersion of feature extraction methods in the identification of plant diseases.

5. Models for Identifying and Categorizing Plant Diseases

In this section, Researchers use various models to classify and detect plant illnesses. These models often rely on deep learning techniques, including CNNs and specialized architectures tailored for plant disease detection. The choice of model influences the accuracy and efficiency of disease identification. This chart generates a pie chart to visualize the distribution of various models used in the context of plant disease identification and categorization. Each model is represented as a segment within the pie chart, and the size of each segment corresponds to the number of studies or applications that have employed that specific model. The chart is aptly titled "Models Used for plant disease Identification and Categorization" to provide context. It serves as an effective visual representation, allowing viewers to discern which models are more commonly utilized in the field. The chart is presented as Figure 12, for easy reference and understanding.

In the study by S. Phani P, et.al. [1], utilized cutting-edge deep learning algorithms such as SSD, YOLOX, and Faster R-CNN models, improved by methods of attention like Efficient Channels Attention, Squeeze & Excitation networks, and Convolutional Block Emphasis Module. The determination of grape leaf disease was done using these types of models. Mahum et al. [4] utilized a previously trained Efficient DenseNet framework, upgrading it with an incremental transition layer to efficiently identify potato leaves into five distinct illness classes. Ashwinkumar et al. [7] demonstrated the efficacy of the MobileNet-based CNN in the identification of plant diseases by using it to diagnose tomato leaf diseases such as early blight, late blight, target spot, and leaf mold. Sharma et al. [8] using CNNs to detect illnesses in potato and rice crops. Deep learning has the ability to automate the identification of plant diseases, as evidenced by the CNN model's better accuracy in disease categorization. Ghosh, S et al. [9] used CNNs in conjunction with more conventional machine learning techniques, such as Support Vector Machine and K-Nearest Neighbors, to classify plants. Their study demonstrates the potential of hybrid methods in this field. Wassan et al. [11] didn't focus on plant disease identification but instead on predicting frost events and their impact on plants, such as leaves and flowers. They employed CNNs for their predictive model, demonstrating the versatility of CNNs in various agricultural applications. Bedi et al. [12] presented a hybrid method that merged convolutional automatic encoders with CNN. This model demonstrates the potential of such hybrid strategies for disease diagnosis and classification; it was particularly created for identifying bacterial spot illness in peach leaves. Chouhan et al.'s study [13] The research incorporated feature extraction methods like scale-invariant and feature transform. These techniques are instrumental in revealing significant characteristics within the image data, which are essential for robust and in-depth analysis. Effective feature extraction is a critical component of plant disease detection, as it enables the model to discern patterns and make accurate classifications. Chelleapandi et al. [14] These comprised DenseNet, MobileNet, MobileNetV2, InceptionV3, InceptionResnetV2, ResNet50, and VGG16, VGG19. To compare its effectiveness with the pre-trained models, the researchers also created and evaluated a specially constructed CNN model. chen et al. [15] adopted a thorough method for identifying diseases, utilizing deep learning and machine learning models. Based on the ideas of meta-learning, they presented the LFM-CNAPS (local feature matching conditional neural adaptive processes) model. Across 26 different plant species, this novel model was able to identify 60 different illnesses. Jadhav et al. [18] used two different CNNs, named AlexNet and GoogleNet, to classify and identify three different diseases in soybean leaves: bacterial blight, brown spot, and frogeye leaf spot. This demonstrates how specific CNN models are used to identify and classify plant diseases. The Sun et al. [24] model was created with the purpose of recognizing and categorizing 26 disease classes that affect 14 distinct plant species. The DM deep learning optimizer, selected for its better performance in comparable situations, served as the foundation of this model. Sujatha et al. [16] For the purpose of classifying diseases, the research combined ML and deep learning (DL) techniques. They employed Random Forest, Stochastic Gradient Descent, and Support Vector Machine (SVM) for machine learning. Conversely, the DL methodologies encompassed Inception- v3, VGG-16, and VGG-19. Abbas et al. [21] used the widely recognized DenseNet model as their main model for tomato leaf disease identification and classification. Furthermore, the study's methodology included a novel integration of C-GAN-produced synthetic pictures. Akshai et al. [22] Their method of identifying and classifying diseases was based on three distinct CNN models. They combined the DenseNet, ResNet, and VGG models to identify different grape leaf diseases. Designed to extract key elements from the grape leaf photos for efficient disease diagnosis, every model offered distinct structural benefits. Umit Atila et al. [23] emphasized the usage of deep learning models extensively. This study relied heavily on the EfficientNet design, namely the B4 and B5 models. Their performance was contrasted with that of other popular CNN models, such as Inception V3, AlexNet, ResNet50, and VGG16. Wang et al. [25] 's main contribution is the SSD architectural improvement known as the DBA_SSD model. The model attains impressive precision in disease identification and categorization by the integration of 1 x 1 convolution, a residual network, and an attention mechanism. DBA_SSD performs better than other popular target detection algorithms such as YOLOv4, YOLOv3, Faster RCNN, and YOLOv4 small. Sharma, P et al. [26]

Two CNN model's performances were compared in the study. While segmented regions-of-interest displaying illness, symptoms were used to train S-CNN, another model, F-CNN, was trained using entire leaf pictures. Karthik R et al. [27] In the study, two deep learning architectures were given. The first model learned features without the requirement for explicit feature engineering by using residual learning with a CNN. By adding an attention mechanism to improve the feature extraction process and enable the model to concentrate on more important regions of the image for disease identification, the second model improved upon the first. Dengshan Li et al. [28] The primary model for classifying and diagnosing pests and diseases that affect rice plants was a proprietary DCNN architecture. When it came to identifying unskilled rice movies, it fared better than other designs such as VGG16, ResNet-50, ResNet-101, and YOLOv3. Ji, M et al. [29] employed the UnitedModel CNN model to identify illnesses of the grape leaf. The paragraph notes that the model was used to categorize three major grape leaf diseases: isariopsis leaf spot, esca, and black rot, but it does not go into great depth on the model's architecture. Nithish, E.K et al. [30] identified five distinct tomato leaf illnessesa CNN model: yellow leaf curl, Septoria leaf spot, early blight, mosaic virus, and bacterial spot. Shrestha, G et al. [32] We out a study with the goal of employing CNNs to identify different diseases in three different plant species: bell pepper, tomato, and potato. This study attempted to identify numerous diseases, including bell pepper bacterial spot, early and late blight in potatoes, target spot, mosaic virus, yellow leaf curl virus, bacterial spot, Septoria leaf spot, leaf mold, early and late blight, and spider mites in tomato plants. The study by Pham, T.N, et al. [33] has the main goal of classifying and identifying illnesses in mango leaves. They used a feed-forward deep neural network model to do this. Marzougui, F et.ai [34] 's Deep learning's potential for plant disease identification is highlighted in this work. Their method fared better than conventional and shallow structural models, proving that deep learning techniques are useful for classifying and diagnosing plant illnesses. In their study, Selvam, L et.ai. [35] created a CNN model for ladyfinger plant diagnosis. The CNN model was created to recognize and classify various ladyfinger leaf states, such as damaged, healthy, and harmed by overuse of fertilizer. M, Turkoglu et al. [36] The study used a combined technique to identify and classify eight different plant diseases. In particular, a number of sophisticated Deep Learning architectures, including GoogLeNet, ResNet-50, ResNet-101, Inception-v3, InceptionResNetv2, and SqueezeNet, were combined with three classifiers: SVM, ELM, and KNN. Too, E.C et al. [37] In order to determine how successful various Deep Learning architectures were at identifying plant diseases, the study team decided to compare and contrast them. VGG- 16, the ResNet series (ResNet-50, ResNet-101, ResNet-152), Inception-V4, and DenseNets-121 were among the architectures assessed in the study. Picon, A. et al. [38] harnessed the capabilities of the ResNet-50 Deep Learning algorithm to craft a mobile application, aiming for the precise identification of plant diseases, particularly focusing on hot spots. M. et al. [39] In the study, a cutting-edge deep learning model named PlantdiseaseNet was presented. This model was created with the purpose of detecting and classifying plant diseases, especially for the demanding agricultural field environment. Zhang et al. [40] presented a unique Deep Learning (DL) model called the multiple Inception-Resnet model that was able to identify yellow rot in wheat plants by combining spectral and spatial data from hyperspectral UAV photos. Kumari, C.U. et al. [41] developed a model to accurately identify illnesses in tomato and cotton leaves by combining neural network technology with image processing techniques. Al Haque et al. [42] entailed using a CNN to identify crop canker, anthracnose, and rot as guava illnesses. M. Francis et al. [43] To detect diseases in the leaves of tomato and apple plants, a CNN was utilized. The main objective of the study was to categorize tomato and apple leaf species into groups of healthy and sick leaves. Sahithya, V. et al. [44] Three different illnesses in lady finger leaves were identified using two categorization techniques: ANN and SVM: powdery mildew, leaf spots, and yellow mosaic vein. M. R. Howlader et al. [45] profound CNN was used to accurately identify and classify illnesses that harm guava leaves, such as rust, whiteflies, and algal leaf spots.

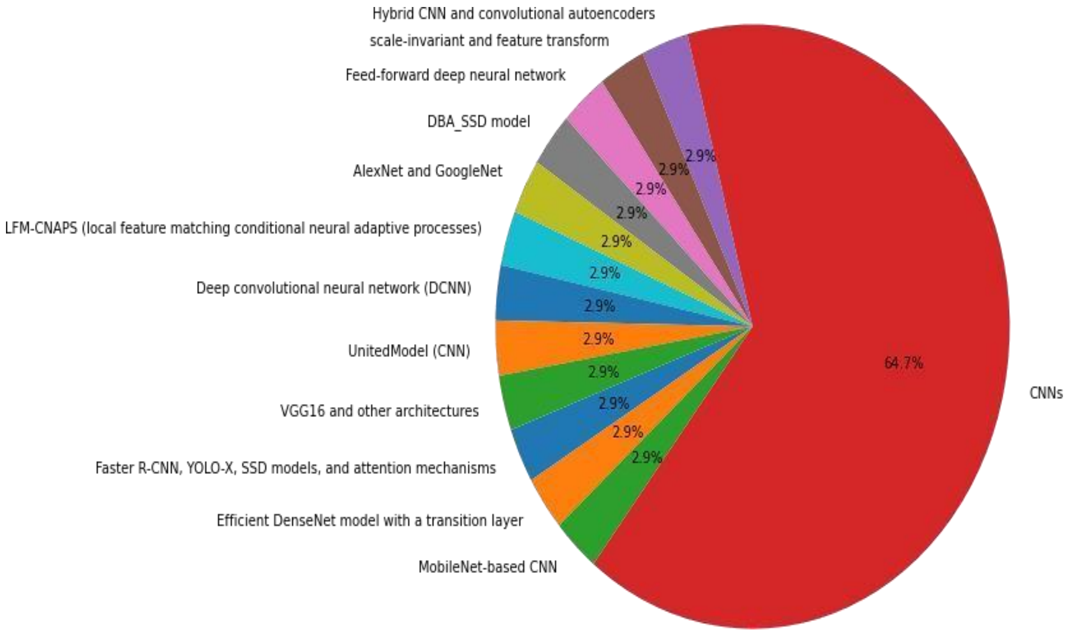

Figure 12: Models Used for Plant Disease Identification and Categorization.

6. Image Quality Improvement Techniques

Techniques to enhance image quality and remove noise are critical in plant disease detection. Researchers employ methods such as brightness adjustment, contrast modification, and image blur. These improvements ensure better image data for analysis. We focus on various filtering techniques aimed at refining the clarity and details of the images. Figure 13, presents a horizontal bar chart depicting the utilization of various image quality improvement techniques in the realm of identifying plant diseases. Each technique is symbolized by a horizontal bar, with the length of each bar indicating the number of research papers employing that specific method. The chart bears the title 'Image Quality Improvement Techniques in Plant Disease Detection,' providing viewers with essential context. To enhance clarity, the y-axis is inverted, positioning the most commonly adopted technique at the chart's top for improved visualization. Additionally, gridlines on the x-axis have been incorporated to enhance readability. Overall, this chart provides a succinct and visual summary of the prevalence of image quality improvement techniques in the realm of identifying plant diseases, facilitating the identification of the most frequently employed methods."

In the study Mahum et al. [4] use of an Efficient DenseNet model suggests an inherent focus on quality feature extraction from the images. Vallabhajosyula et al. [6] During the initial image processing phase, special attention was given to optimizing the photos' contrast and brightness. This enhancement significantly improved the overall quality of the dataset. A bilateral filter was used by Ashwinkumar et al. [7] during pre-processing to enhance image quality by eliminating noise from plant leaf images. Malathy et al. [10] the importance of image quality in plant disease detection. Analytical approaches, such as image restoration, were used to elevate the quality, clarity, and reliability of the images used in the study. Chowdhury et al. [19] The application of resizing and normalization during pre-processing enhances image quality and prepares the data for further analysis, highlighting the importance of image quality in disease detection. M. R. Howlader et al. [45] used the Rectified Linear Unit (ReLU) activation function to enhance the overall appearance of the images. F(N) = Max (0, N) is the representation of the operation, where N is the total amount of neurons.

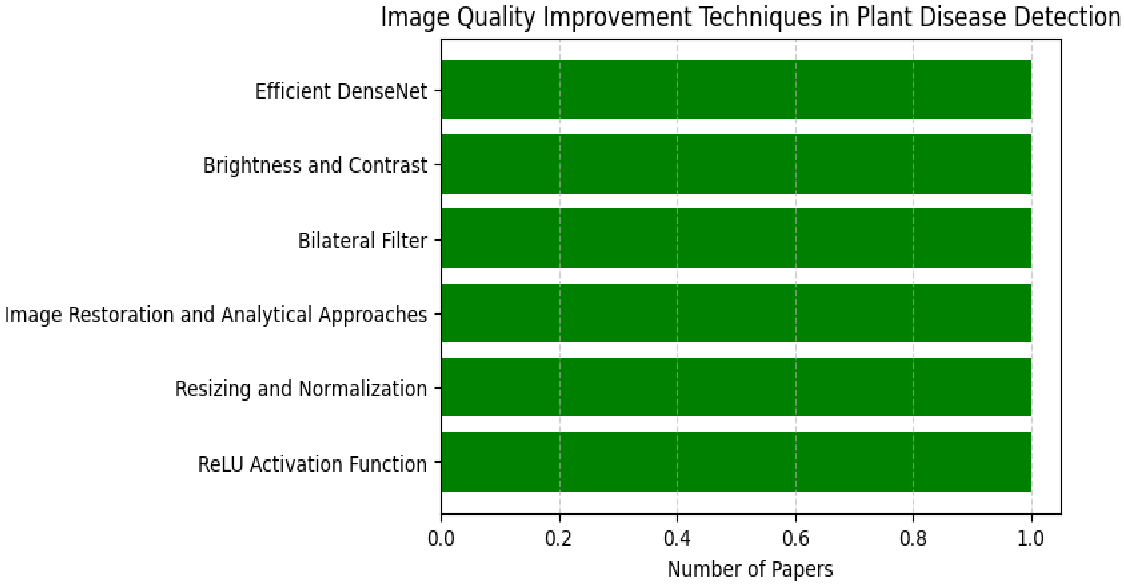

Figure 13: Overview of Image Quality Improvement Techniques.

7. Strategies for Reducing Overfitting in Plant Disease Detection

Overfitting is a common challenge in machine learning. Researchers use strategies like data augmentation, regularization techniques, and early stopping to prevent models from fitting the training data too closely, thus improving the model's generalization. we address the issue of overfitting, which occurs when a model produces significantly different accuracy values for training and testing datasets. Given below this chart have to illustrate the distribution of strategies employed to reduce overfitting in the context of plant disease detection. Each literature is represented as a slice of the pie, and the size of each slice corresponds to the number of research papers that have utilized that particular Model. The chart is titled 'Strategies for Reducing Overfitting in Plant Disease Detection,' providing essential context. By displaying this information in a pie chart, viewers can readily discern the prevalence of different overfitting reduction strategies within the field. The chart is presented as Figure14, to reference it within the document.

Pandian et al. [5] To prevent the model from overfitting, data augmentation techniques were implemented. These techniques help ensure that the model generalizes well to new data and doesn't become excessively specialized to the training set. Vallabhajosyula et al. [6] To ensure our model's robustness and prevent it from becoming too specialized to the training data, we successfully decreased overfitting by using strategies for data augmentation. Bedi et al. [12] Addressing the pervasive issue of overfitting in deep learning models, the researchers implemented "early halting." By setting a patience value to 5, they efficiently halted the training once the model started to overfit, ensuring its generalization capability when applied to unseen data. The paper [15] emphasizes the use of forward propagation within the LFM-CNAPS model to address potential overfitting issues. This strategy ensures that the model remains generalizable across varied disease classes and plant species. Chowdhury et al. [19] The study employed Global Average Pooling (GAP) as an approach to minimize overfitting. This technique is effective in preventing models from learning noise in the data and improving generalization. Lijo, J. et al. [20] utilization of data augmentation techniques, as mentioned above, served as a strategy for reducing overfitting. The investigators sought to enhance the generalization and effectiveness of the model's predictions by adding changes to the dataset. Wang et al. [25]'s DBA_SSD model, with its incorporation of 1 × 1 convolution, is strategically designed to counteract overfitting. Overfitting can lead to poor model generalization on unseen data, and measures such as this ensure that the model remains robust and accurate across diverse data inputs. Ji, M et al. [29] To mitigate model overfitting, the researchers implemented several measures. Those included dropout techniques,

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. X No. Y (20--)

data augmentation schemes, and the implementation of an early stop mechanism. These techniques are meant to keep the model from learning the training set by heart and improve its ability to generalize to fresh, untrained data. Nithish, E.K et al. [30] Various data augmentation strategies were employed in the research to mitigate model overfitting. These strategies help the model generalize better to unseen data by exposing it to a wider rangeof augmented examples during training. Chen, J et al. [48] data augmentation techniques served a dual purpose. While they expanded the dataset by introducing variations of the original images, they also acted as an effective strategy to minimize overfitting, ensuring that the model remains generalized and versatile. M. Francis et al. [43] incorporated a dropout value of 0.25 in the CNN architecture to mitigate overfitting, ensuring the model's generalization to new data. M. R. Howlader et al. [45] The study employed various data augmentation strategies and the Rectified Linear Unit (ReLU) activation function to address concerns about overfitting. These methods were critical in enhancing the performance of the model.

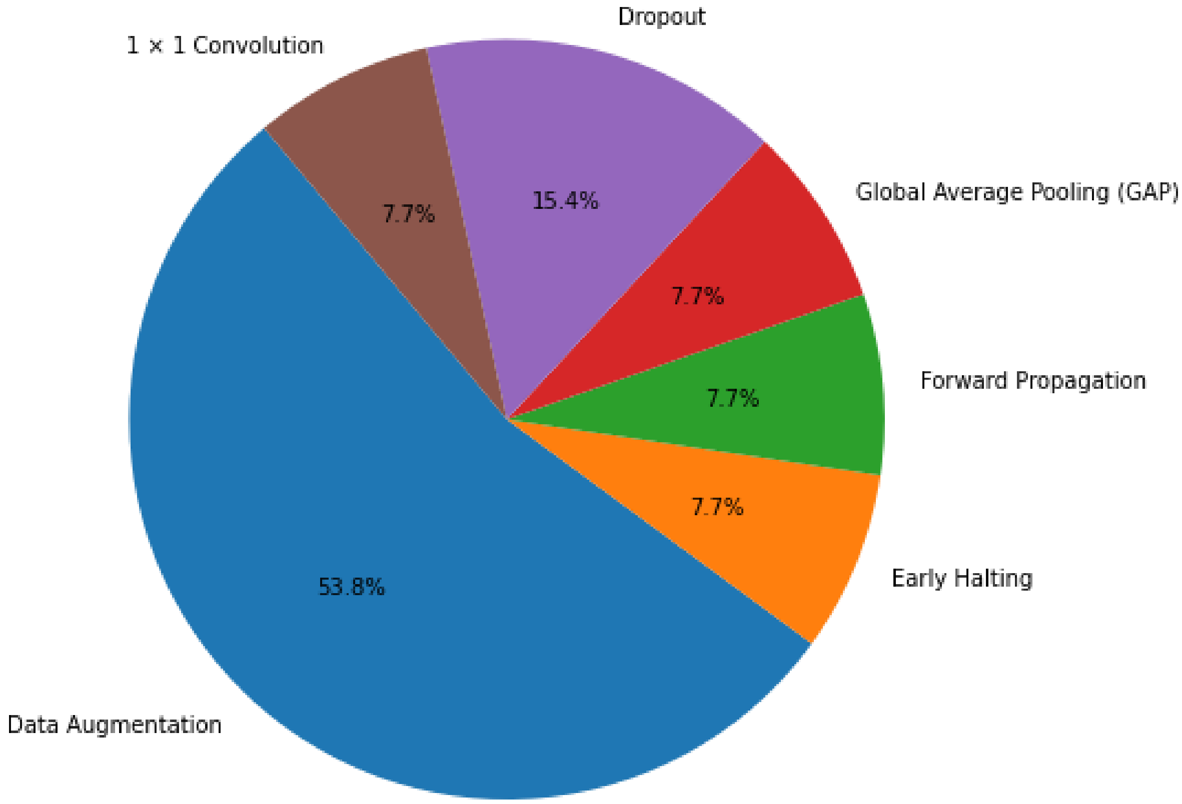

Figure 14: Distribution of Strategies for Reducing Overfitting.

8. Plant Species and Disease Classes in Evaluated Studies

This section discusses the various plant species and categories of illnesses studied in the literature. Different studies concentrate on particular plant species, like tomato, mango, and ladyfinger, while evaluating various disease classes for detection. Our attention is directed toward the disease categories identified within the particular plant species that served as the foundation for the studies under review. The chart, presented as Figure 15, illustrates a histogram showcasing the distribution of plant species and disease classes within a set of evaluated studies. This visual representation compiles data related to different plant species and their corresponding disease classes. By tallying the occurrences of distinct categories representing various plant species and disease classes, the chart effectively conveys their prevalence. The y-axis of the histogram denotes the counts, while the x-axis enumerates the categories, facilitating a clear understanding of which plant species and disease classes receive the most attention in these studies. To enhance readability, the x- axis labels are thoughtfully rotated. The chart is aptly titled "Distribution of Plant Species and Disease Classes in Evaluated Studies." Overall, this histogram offers a comprehensive overview of the distribution of these categories across the analyzed studies.

In the study by S. Phani P, et.al. [1], mainly concentrates on identifying grape leaf diseases, with particular classes covering Leaf Spot, Black Rot, Esca measles, and healthy circumstances. Mahum et al. [4] The study zeroes in on potato leaves, classifying them into five distinct categories: Potatoes Leaves Rolls (PLR), Potato Verticillium wilt (PVw), Potatoes Late Blight (PLB), Potato Early Blight (PEB), and Potatoes Healthy (PH). Vallabhajosyulaet al. [6] Our approach centered on developing a deep ensemble neural network capable of diagnosing 38 differentdisease classes across 14 plant species. Malathy et al. [10] focused on apple photos, identifying illnesses such sooty spot, powdery mildew, and bittersweet rot in particular. Their emphasis on these diseases underlines the significance of timely detection and intervention to prevent loss in apple cultivation. Bedi et al. [12] concentratedon peach leaves, aiming to identify bacterial spot disease, a prevalent issue in peach cultivation. The LFM-CNAPS model, as proposed by Chen et al. [15], was used to identify an extensive range of 60 diseases across 26 differentplant species, showcasing its versatility and breadth in plant disease diagnosis. Jadhav et al.'s [18] research primarily focuses on soybean, with the study aiming to identify three specific diseases that affect this plant species(bacterial blight, brown spot, and frogeye leaf spot). This comprehensive categorization facilitates the assessmentof the efficacy and precision of the proposed models. Sun et al. [24] model was geared towards recognizing 26 distinct disease classes that affect the foliage of fourteen distinct plant species. This showcases the comprehensivenature of the study in covering a wide range of diseases across multiple crops. Abbas et al [21] used a technique to pinpoint nine distinct illnesses in tomato leaves. These provide a wide range of tomato leaf disease detection, such as yellow leaf curl virus, bacterial spot, Septoria leaf spot, two-spotted spider mite, target spot, early blight, leaf mold, late blight, and mosaic viruses. Akshai et al. [22] primary focus was on grape leaves, specifically identifying three prevalent diseases: black rot, leaf blight, and esca. The aim was to offer a comprehensive disease detection mechanism for grape leaves, considering the economic and agricultural significance of this crop. Wanget al. [25] encompasses various plant species and disease severity levels, covering a comprehensive range of 15 subcategories. This wide coverage ensures that the model is tested against a diverse array of diseases, enhancing its reliability. Karthik R et al. [27] specifically concentrated on tomato plants, examining three key disease classes Leaf mold, late blight, and early blight are all included in the Plant Village Dataset. Dengshan Li et al. [28] 's the primary focus was on rice plants, with a particular emphasis on rice brown spot, rice stem borer symptoms, and rice sheath blight. A particular emphasis was placed on heavily infested instances to enhance grain production by controlling these issues. Ji, M et al. [29] specifically focused on grape leaf diseases, with the classification and diagnosis targeting three main afflictions: black rot, esca, and isariopsis leaf spot. Nithish, E.K et al. [30] concentrated on tomato leaf diseases with a specific goal of diagnosing five diseases: bacterium spot, yellow leaf curl, early blight, septoria leaf spot, and mosaic virus. Shrestha, G et al. [32] covered three plant species: potato, tomato, and bell pepper. Within these species, the research delved into the identification of twelve distinct disease classes, including two for potato, nine for tomato, and one for bell pepper. This comprehensive analysis encompassed a diverse range of diseases affecting the leaves of these plant species. Too, E.C et al. [37] covered an extensive range of plants, including the following: potato, peach, raspberry, apple, bell pepper, blueberry, cherry, corn, orange, grape, soybean, strawberry, tomato, and squash. This expansive coverage ensures a comprehensive understanding of disease detection across different plant species. Picon, A. et al. [38] predominantly focused on diseases affecting wheat plants, emphasizing the importance of addressing agricultural challenges in staple crops. M. et al. [39] covered a wide range of plant species, containing tomato, sugar beets, apple, bell pepper, cherry, grapes, onion, peach, potato, plum, or wheat crops, emphasizing the diversity of plants studied. Kumari, C.U. et al. [41] The suggested technique effectively identified illnesses in tomato and cotton leaves, such as Septoria leaf spot, leaf mold, and targeting spot in tomatoes.

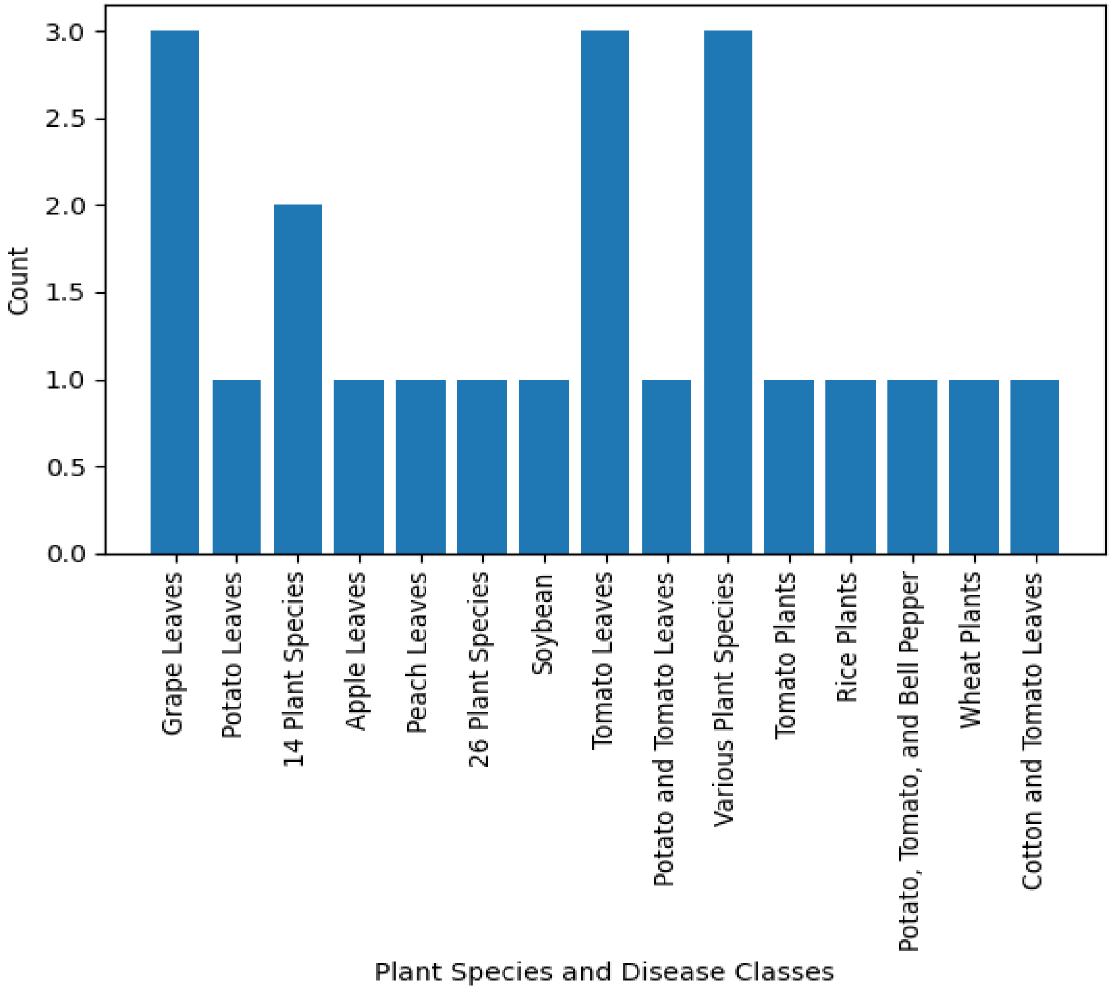

Figure 15: Distributions of Plant Species and Disease Classes in Evaluated Studies.

9. Accuracy in Plant Disease Identification

An important part of the research is determining how accurate the disease identification models are. Scholars frequently present accuracy measures to assess how well their algorithms classify and identify plant diseases. The review of current methods for the identification and categorization of plant diseases, with an emphasis on their accuracy, is the main focus of this phase. The chart, presented as Figure 16, is a pie chart designed to visualize the distribution of studies based on their accuracy levels within the field of plant disease identification. The chart effectively classifies these studies into four primary groups: "High Accuracy," "Average Accuracy," "Low Accuracy," and "Other."

In the "**HighAccuracy**" category, you'll find studies that have achieved impressive accuracy levels, such as those involving CNN models with ResNet architecture, deep ensemble neural networks, DenseNet, DenseNets architecture, MobileNet-based CNN, and a proposed method. These studies have demonstrated accuracy levels that are in proximity to, or surpass, the 97% mark.

The "**Average Accuracy**" category encompasses studies that exhibit accuracy levels within the range of 87.4% to 97%. These studies employ a variety of models and techniques, including potato leaf disease detection, high accuracy in identifying apple, rice, and potato diseases, C-GAN models, GoogleNet, VGG- 16, EfficientNet models, DBA_SSD models, AlexNet, UnitedModel, and custom CNN models.

For the "**Low Accuracy**" category, it accommodates studies characterized by relatively lower accuracy levels, such as those featuring CNN models with an 87% accuracy rate, test accuracy results at 88.8%, a training accuracy of 91.32% coupled with testing accuracy of 85.45%, custom DCNN systems designed for efficiency, RF-classifiers with a 77% accuracy, and studies with balanced accuracy levels of 87%.

The "**Other**" category encompasses studies that don't neatly fit into the aforementioned accuracy groupings. These studies involve multiple accuracies for different diseases, SVM and ANN models exhibiting varying accuracy rates, the use of PlantDiseaseNet techniques yielding a classification accuracy of 93.67%, DCNN models exhibiting improved accuracy, as well as studies demonstrating high accuracy in detecting bacterial leaf spots and tomato diseases. Some studies in this category present various accuracy levels where specific values are not provided.

In the study by S. Phani P, et.al. [1], The YOLO-X and SSD models, which were combined with attention approaches, outperformed the other examined models in terms of accuracy and real-time performance. The YOLO-X models proved particularly efficient, excelling in detecting smaller objects and those obscured by background interference, thus offering robust and accurate results for real-time grape disease diagnosis in vineyards. Singh, G, et.al. [2] use of deep learning techniques, including base learning and transfer learning, the study aimed to identify and classify citrus diseases. The proposed CNN model, particularly when employing the ResNet architecture, delivered an outstanding accuracy rate of 99.62%. This remarkable achievement surpassed the performance of other contemporary models, highlighting the efficacy of the proposed approach in accurately identifying and categorizing plant diseases, particularly those affecting potato leaves. Mahum et al. [4] The implemented algorithm showcased a remarkable accuracy of 97.2%, outpacing several existing models in the domain of potato leaf disease detection. Vallabhajosyula et al. [6] Using the PlantVillage dataset, we assessed the performance of our deep ensemble neural network technique and obtained an astounding accuracy rate of 99.99%. This underscores the effectiveness of our approach in accurately diagnosing plant diseases. Through the implementation of the proposed MobileNet-based CNN, Ashwinkumar et al. [7] shown a remarkable 98.7% accuracy in recognizing and classifying different illnesses of tomato leaves, highlighting the usefulness of the model for diagnosing plant illnesses. In terms of rice and potato disease identification, the CNN model put forth by Sharma et al. [8] shown impressive accuracy rates of 99.58% and 97.66%, respectively. These high levels of accuracy highlight how well their deep learning method works for diagnosing plant diseases. Ghosh, S et al. [9] the study showcases impressive classification accuracy, with the CNN + K-Nearest Neighbors approach achieving high accuracy levels, such as 99.5% in Flavia, 97.4% in LeafSnap, and 80.02% in MalayaKew datasets, underscoring the effectiveness of their hybrid model. Malathy et al. [10] The effectiveness of the suggested CNNmodel was demonstrated when it was able to identify certain diseases in apple pictures with an astounding 97% accuracy rate. This high accuracy emphasizes the potential and efficacy of the CNN model in plant disease detection. Bedi et al. [12] approach achieved remarkable accuracy rates. For the testing dataset, the accuracy wasan impressive 98.38%, while the training dataset showcased an even higher accuracy of 99.35%. This high level of accuracy not only validated their method but also emphasized its superiority compared to other available techniques in the realm of plant disease detection. The DenseNet model, as highlighted by Chelleapandi et al. [14], reached a remarkable 99% accuracy rate in classifying and diagnosing plant diseases. This highlights the deep learning models' great accuracy potential, especially when it comes to plant disease identification. Employing the LFM-CNAPS model rooted in meta-learning, Chen et al. [15] were able to identify diseases in all examined plant species with an impressive accuracy rate of 93.09%. This accomplishment highlights the model's effectiveness and meta-learning's promise for plant disease identification. Jadhav et al.'s [18] For the purpose of diagnosing illnesses from soybean leaves, the AlexNet model obtained an accuracy of 98.75%, whereas the GoogleNet model obtained an accuracy of 96.25%. These high accuracy rates highlight how well the suggested models identify diseases. Sun et al. [24] proposed model employing the The DM deep learning optimizer producedresults with an astounding 97% accuracy. This demonstrates the high degree of reliability of the model in detecting and classifying plant illnesses and attests to the effectiveness of the DM optimizer in this regard. Sujatha et al.
[16] In terms of diagnosing illnesses in citrus plants, DL models performed better than ML models. Whereas DL models, such as Inception-v3, VGG-16, and VGG-19, produced much better illness detection accuracy levels,

ranging from 87.4% to 89.5%, ML models acquired accuracies ranging from 76.8% to 87%. This demonstrates the DL systems' improved performance in the particular setting of citrus plant illnesses identification. Abbas et al. [21] presented remarkable results. The C-GAN model, responsible for generating synthetic images, achieved an impressive accuracy of 99.51%. Subsequently, when the DenseNet model was applied for disease classification, the accuracy rates achieved were 98.65% for five classes, 97.11% for ten classes, and 98.65% for seven classes of tomato leaf images, reinforcing the model's proficiency in disease detection and classification. Akshai et al. [22] among the three CNN models employed, DenseNet emerged as the most effective, achieving a remarkable accuracy rate of 98.27% in detecting the specified diseases in grape leaves. Such an impressive accuracy rate not only affirms the model's efficiency but also underscores the potential of CNNs in plant disease identification. Umit Atila et al. [23] On the original dataset, the B5 model of EfficientNet demonstrated an accuracy of 99.91%, whereas the B4 model demonstrated an accuracy of 99.97% on the enhanced dataset. This demonstrates the EfficientNet architecture's potential for classifying plant leaf diseases. Wang et al. [25] The proposed DBA_SSD model attains an impressive accuracy of 92.20% in detecting and classifying plant diseases.Such high accuracy rates underscore the model's potential in real- world applications and its reliability in diseasedetection. Sharma, P et al. [26] Notable outcomes were found when comparing the F-CNN and S-CNN models. On one dataset, the F-CNN model's accuracy was 96.3%; however, when tested on separate datasets, it fell to 42.3%. In contrast, the S-CNN model consistently performed well, maintaining an accuracy rate of 98.6% on independent data. This highlights the potential advantages of using segmented and annotated images for training deep learning models in the context of plant disease detection. Karthik R et al. [27] achieved notable success in their disease identification efforts. By incorporating the attention mechanism into their model, they used a 5-foldcross- validation setup and got an astounding overall accuracy of 98% on the validation sets. The efficacy of theirmethodology and the possible advantages of using attention mechanisms in plant disease detection models are shown by this high accuracy. Dengshan Li et al. [28] pecific accuracy values were not provided in the paragraph,it is claimed that the bespoke DCNN system showed promising detection results, especially for indications of ricestem borer and rice sheath blight. The system also achieved a fast detection speed of approximately 0.1 seconds per frame, indicating its efficiency in processing video data. Ji, M et al. [29] A test accuracy of 98.57% and a validation accuracy of 99.17% were attained by the suggested UnitedModel. These impressive accuracy rates show how well deep learning works for correctly diagnosing grape diseases, including black rot, esca, and isariopsis leaf spot. Nithish, E.K et al. [30] The suggested CNN model successfully identified these five diseasesfrom tomato leaf photos with an astounding 97% accuracy rate, demonstrating the algorithm's efficacy in diseasedetection. The outcome of the research by Shrestha, G et al. [32] demonstrated a noteworthy 88.8% test accuracyin identifying diseases impacting the leaves of plants such as bell pepper, potato, and tomato. This accuracy statistic highlights how well the CNN model classifies and recognizes a range of plant illnesses in diverse plant species. The deep neural network model proposed in the research by Pham, T.N, et.ai. [33] displayed a notable level of performance. In the identification of mango leaf diseases, it had a testing accuracy of 85.45% and a training accuracy of 91.32%. The efficacy of the algorithm in precisely categorizing various disease categories inmango leaves is demonstrated by these accuracy metrics. M, Turkoglu et al. [36] he combination of the ResNet- 50 model with the SVM classifier demonstrated remarkable efficacy, achieving an F1 score of 97.14% and an overall accuracy of 97.86%. Too, E.C et al. [37] all the architectures assessed, the DenseNets architecture stood out for its exceptional performance, achieving a remarkable testing accuracy of 99.75% in plant disease identification. This underscores the potential and superiority of the DenseNets model in this domain. Picon, A. etal. [38] When the model's performance was measured using Balanced Accuracy, it demonstrated a remarkable 87% accuracy rate, demonstrating the model's dependability in the diagnosis of diseases affecting wheat plants.

M. et al. [39] PlantDiseaseNet methods showed remarkable efficacy, with a 93.67% classification accuracy (CA). This demonstrates how well the model can recognize and categorize plant diseases in actual agricultural situations. X., Zhang et al. [40] DCNN model exhibited a significantly improved accuracy of 85% compared to the RF-classifier, which achieved an accuracy of 77%. This illustrates how successful deep learning methods are specifically DCNN, for enhancing disease detection in wheat plants within a real agricultural environment. Kumari, C.U. et al. [41] The model demonstrated remarkable accuracy rates: it identified two different types of tomato illnesses from tomato leaves with an astounding 100% accuracy, 90% accuracy for bacterial leaf spots, and 80% accuracy for target spots on cotton leaves. These outcomes show how well the suggested approach can correctly identify and classify diseases of plants. Al Haque et al. [42] The proprietary CNN model demonstrated the efficacy of the methodology in illness detection, identifying guava diseases with an astounding detection rate of 95.61%. M. Francis et al. [43] In the datasets containing apple and tomato leaf samples, the CNN model had an 87% accuracy rate in identifying illnesses, reflecting its effectiveness in plant disease identification. Sahithya, V. et al. [44] models successfully diagnosed these diseases, with SVM achieving an accuracy of 85% in the presence of noisy images and 92% without noise, while ANN achieved 97% accuracy with noise and 98% without noise in the image. M. R. Howlader et al. [45] The suggested approach demonstrated its proficiency in identifying illnesses including fruit rot, anthracnose, and fruit canker on guava leaves with an amazing accuracy rate of 98.74%.

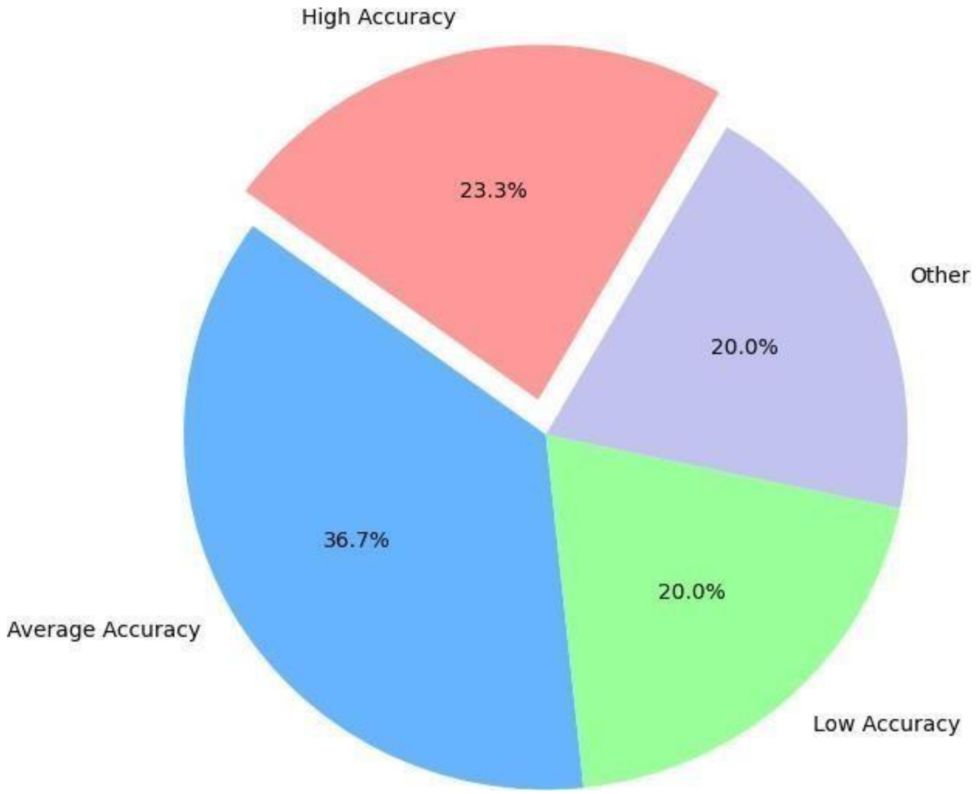

Figure 16: Distribution of Accuracy Levels in Plant Disease Identification.

Overall, the pie chart serves as a visual summary that elucidates the distribution of these studies across distinct accuracy categories. This clear visualization makes it simple to comprehend the prevalence of different accuracy levels within the domain of plant disease identification studies. The chart is appropriately titled "Distribution of Studies by Accuracy Level in Plant Disease Identification" to provide added clarity and context.

4. Comparative Analysis

We conduct a thorough evaluation of the different disease detection and classification models used in agriculture in the comparative analysis section. The primary objective is to scrutinize the performance and capabilities of these models in accurately identifying and categorizing different plant diseases. Each model's technical underpinnings and key techniques are thoroughly examined, shedding light on their distinctive approaches. Additionally, the datasets utilized in these studies are considered, evaluating their representativeness and suitability for real-world applications. A wide spectrum of disease types is investigated, ensuring the breadth of the research's applicability. The accuracy rates achieved by these models are scrutinized, providing valuable insights into their efficacy. The primary goal of this section is to assist researchers and practitioners in choosing the best method for their unique needs related to agricultural disease detection and categorization by outlining the advantages and disadvantages of each model. Summary of existing review table 1 displays the comparison analysis.

Summary of existing review table 1.

Ref No.	**Disease Identify/Classify Model Used**	**Key Techniques Used**	**Used dataset**	**Disease Types Investigated**	**Accuracy (%)**	**Primary Focus Highlight**
[1]	R-CNN, YOLO-X, and SSD with Attention Mechanisms that operate faster	Deep Learning	Plant Village Dataset (Grape Leaves)	Black Rot, Esca measles, Leaf spot, Healthy	High Precision and Real-time Performance	Efficient Disease Detection with Attention Mechanisms
[2]	CNN (ResNet)	CNN-Based Identification	'Plant Village' Dataset (Potato Leaves)	Potato Diseases	Impressive Accuracy	Potato Disease Identification with CNN (Focus: Potato Leaves)
[4]	Efficient DenseNet	Deep Learning with Pre-processing	'The Plant Village'Datase t (Potato Leaves)	Potato Diseases	97.2% Accuracy	Improved Potato Leaf Disease Classification (Focus: Potato Leaves)
[5]	Conv-5 DCNN	Data Augmentation and Deep Learning	Public Datasets	Various Plant Species	98.41% Accuracy	High Accuracy Disease Classification with Conv-5 DCNN (Focus: Various Plant Species)
[6]	Deep Ensemble Neural Network	Data Augmentation and Ensemble Models	Plant Village Dataset (Multiple Plant Species)	Plant Diseases	99.99% Accuracy	Enhanced Plant Disease Identification with Deep Ensemble Neural Network (Focus: Multiple Plant Species)
[7]	MobileNet- based CNN	Image Pre-processing and CNN	Tomato Leaf Diseases	Tomato Diseases	98.7% Accuracy	Tomato Disease Identification Using MobileNet-based CNN (Focus: Tomato Leaves)
[8]	CNN (Comparative Study)	CNN-Based Diagnosis	Large Datasets (Rice and Potato Leaves)	Rice and Potato Diseases	High Accuracy	Comparative CNN Analysis for Rice and Potato Disease Identification (Focus: Rice and Potato Leaves)

[9]	CNN + SVM, CNN + kNN	Hybrid Models	Diverse Datasets (LeafSnap, Flavia, MalayaKew)	Plant Classification	Impressive Accuracy	Hybrid Models for Accurate Plant Classification (Focus: Plant Classification)
[10]	CNN	Image Pre-processing and CNN	Image Dataset (Apples)	Apple Diseases	97% Accuracy	Apple Disease Identification with CNN (Focus: Apple Diseases)
[11]	CNN (Convolutional Neural Networks)	CNN-Based Frost Prediction	Historical Weather Data	Frost Prediction	98.6% Accuracy	Frost Event Prediction with CNN (Focus: Frost Prediction)
[12]	Hybrid CNN and Convolutional Autoencoders	Hybrid CNN	Peach Leaf Diseases	Bacterial Spot Detection	High Accuracy	Peach Leaf Disease Detection Using Hybrid Models (Focus: Peach Leaves)
[13]	RESNET18, LFM-CNAPS	Feature Extraction and Meta-learning	Leaf Images (Various Plant Species)	Plant Disease Classification	93.09% Accuracy	Meta-learning Approach for Comprehensive Plant Disease Classification (Focus: Plant Classification)
[17]	CNN	Image Pre-processing and Feature Extraction	New Plant Diseases Dataset (Maize, Grapes, Rice)	Disease Features	Not Specified	Pre-processing Images and Extracting Features to Identify Diseases (With an Emphasis on Disease Features)
[18]	AlexNet, GoogleNet	CNN-Based Disease Recognition	Soybean Leaf Images	Frogeye Leaf Spot, Brown Spot, Bacterial Blight	High Accuracy	Disease Identification with AlexNet and GoogleNet (Focus: Soybean Leaves)
[19]	Not specified but mentioned EfficientNet and U-net	Data preprocessing, normalization, data augmentation	PlantVillage dataset (Tomato Leaves)	Various tomato leaf diseases	Not specified	Utilization of EfficientNet and U- net for Tomato Disease Detection and (GAP) used Overfitting Reduction
[20]	Not specified	Data scaling, data augmentation	Plant Village dataset (Various Plant Leaves)	Various bacterial and fungal plant illnesses	Not specified	Data Augmentation for Diversity and Overfitting Reduction

[24]	DM (Discount Momentum) deep learning optimizer	Disease classification, multiple crop types	PlantVillage dataset (Multiple Crops)	26 disease classes among 14 crop types	97%	Efficient Disease Classification with DM Deep Learning Optimizer
[16]	ML (SVM, Stochastic Gradient Descent, Random Forest) and DL (Inception- v3, VGG-16, VGG-19)	Comparison of ML and DL models	Citrus leaves from citrus research center	Citrus disease classification	ML models: 76.8% - 87%; DL models: 87.4% - 89.5%	DL Models Outperform ML Models for Citrus Disease Detection
[21]	DenseNet model and C- GAN for synthetic images	Data augmentation, synthetic image generation	PlantVillage dataset (Tomato Leaves)	Multiple tomato leaf diseases	Various (e.g., 98.65% for 5 classes)	Effective Identification and Classification of Tomato Leaf Diseases
[22]	VGG, DenseNet, ResNet	Data augmentation (rotation, shifting, zooming)	PlantVillage dataset (Grape Leaves)	Black rot, leaf blight, esca	98.27%	Efficacy of CNN-Based Models for Grape Disease Detection
[23]	EfficientNet (B4, B5), AlexNet, ResNet50, VGG16, Inception V3	Comparison of EfficientNet with other CNN models	PlantVillage dataset (Multiple Plant Species)	Multiple plant diseases	EfficientNet: 99.91% - 99.97%	Efficiency of EfficientNet for Plant Disease Classification
[25]	DBA_SSD	Enhanced SSD architecture, attention mechanism	PlantVillage dataset (Various Plant Leaves)	Plant diseases (15 subcategories)	92.20%	Effective Object Detection Algorithm for Plant Diseases
[26]	F-CNN and S-CNN (Segmented images)	Data augmentatin (brightness adjustment, contrast, blur)	Various datasets (Tomato Leaves)	Tomato leaf diseases	F-CNN: 96.3%, S-CNN: 98.6%	Benefit of Using Segmented Images for Disease Detection
[27]	CNN models with residual learning and attention mechanism	Data augmentation	Plant Village dataset (Tomato Leaves)	Tomato leaf diseases	98%	Effective Disease Detection in Tomato Plants
[28]	Personalized architecture for deep convolutional neural networks (DCNNs)	Data augmentation	Rice plant images and videos	Rice diseases and pests	Not specified	Effective Detection of Rice Diseases and Pests in Videos
[29]	Convolutional Neural Network (UnitedModel)	Data augmentation (rotation, zooming, flipping, shearing, color alterations)	Not specified	Diseases of grapes: esca, black rot, and isariopsis leaf spot	Test: 98.57%, Validation: 99.17%	Successful Diagnosis of Grape Leaf Diseases with Deep Learning

[30]	Convolutional Neural Network (CNN)	Data augmentation (RandomResized Crop, Random Rotation)	Not specified (Tomato Leaves)	Five tomato leaf diseases	97%	Data Augmentation for Tomato Disease Detection
[48]	Not specified	Data augmentation (rotation, flipping, scaling, translation)	Fujian Institute of Subtropical Botany dataset (Rice and Maize Leaves)	Various rice and maize diseases	Not specified	Expanding Real-World Dataset for Disease Identification
[32]	Convolutional Neural Network (CNN)	Disease classification, multiple plant species	Potato, Tomato, Bell Pepper Leaves	Multiple diseases in potato, tomato, and bell pepper plants	88.8%	Disease Identification in Multiple Plant Species
[33]	Feed-forward deep neural network	Preprocessing (downscaling, contrast enhancement)	Giang Province, Vietnam dataset (Mango Leaves)	Powdery mildew, anthracnose, gall midge, healthy leaves	Training: 91.32%, Testing: 85.45%	Effective Mango Disease Identification
[34]	Not specified	Data augmentation (horizontal flipping, rotations, shifts)	Not specified	Not specified	Not specified	Deep Learning for Disease Detection
[35]	Convolutional Neural Network (CNN)	Data augmentation (rotation, flipping, shearing, zooming, shifts)	Tiruvannamalai district dataset (Lady Finger Leaves)	Lady finger leaves with diseases and excessive fertilizer use	Not specified	Robust Model for Ladyfinger Leaf Identification
[36]	SVM, Extreme Learning Machine, K-Nearest Neighbor, Deep Learning models	Multiclass classification, various Deep Learning architectures	Actual field dataset (cherry, walnut, peach, and apricot)	Various plant diseases	F1: 97.14, Accuracy: 97.86%	Comparative Evaluation of Deep Learning Models for the Identification of Plant Diseases
[37]	Not specified	Evaluation of multiple Deep Learning architectures	PlantVillage dataset (Multiple Plant Species)	Multiple plant diseases	Testing accuracy: 99.75%	Determining the Most Effective Deep Learning Architecture for Disease Detection
[38]	ResNet-50	Extension of a deep residual neural network	Real-world dataset (Wheat plants)	Wheat plant diseases	87%	Developing a Mobile Application for Plant Disease Detection

[39]	PlantDiseaseNet	Deep Learning model for agricultural settings	Real-world datasets (Multiple Plant Species)	Multiple plant diseases	93.67%	Disease Detection in Real Agricultural Settings
[40]	Multiple Inception- Resnet model	Hyperspectral UAV images for disease detection	Real-world datasets (Wheat Leaves)	Yellow rust in wheat plants	85%	Improved Disease Detection with Hyperspectral Images
[41]	Image processing and neural networks	Feature extraction (GLCM)	Cotton and Tomato leaves	Cotton (target spot, bacterial leaf spot), Tomato (leaf mold, septoria leaf spot)	Cotton: 80% -90%, Tomato: 100%	Successful Disease Identification in Cotton and Tomato Leaves
[42]	Custom CNN model	Data augmentation (horizontal flipping, zooming, shifting, rotation, shearing)	Real-world dataset (Guava Leaves)	Guava diseases (fruit canker, anthracnose, fruit rot)	95.61%	Effective Guava Disease Detection
[43]	CNN	Preprocessing (resizing, cropping)	Not specified (Apple and Tomato Leaves)	Apple and tomato leaf diseases	87%	Effective Disease Identification in Apple and Tomato Leaves
[44]	SVM, ANN	Feature extraction (GLCM), multiple classification techniques	Real-world dataset (Lady's Finger Leaves)	Lady's finger leaf diseases (powdery mildew, leaf spots, yellow mosaic vein)	SVM: 85% - 92%, ANN: 97% - 98%	Disease Detection in Lady's Finger Leaves
[45]	Deep CNN	Preprocessing (resizing)	Not specified (Guava Leaves)	Guava diseases (algal leaf spots, rust, whiteflies)	98.74%	Effective Detection of Guava Leaf Diseases

5. Conclusion

In conclusion, an extensive analysis of the use of image processing methods for the identification of plant leaf diseases is given in this review study. Plant diseases are still a major danger to agricultural productivity, hence there is a growing need for precise and effective detection techniques. One effective method for overcoming this difficulty is the combination of deep learning, machine learning, and image processing techniques. Through an extensive literature review, we have explored various aspects of disease detection in plants, from model selection and feature extraction strategies to data sources and pre-processing approaches. We have also delved into the methods for improving image quality and addressing overfitting issues, ultimately aiming to enhance the accuracy of disease identification. The studies reviewed in this paper have showcased the breadth and depth of research in this field. Researchers have leveraged diverse data sources, enhanced the quality of image data, and employed a range of feature extraction methods to identify diseases in various plant species. Machine learning and deep learning models have played a pivotal role, achieving impressive accuracy rates in disease classification. Additionally, image quality enhancement techniques and strategies to mitigate overfitting have contributed to the robustness of these models.

6. Future scope

Research and development opportunities abound in the field of plant leaf disease detection utilizing image processing techniques. As technology continues to advance, the following areas are worth exploring:

1. **Real-time Detection: -** It would be very helpful to build real-time illness detection tools for use in the field. This could involve the integration of image processing techniques with IoT (Internet of Things) devices to provide farmers with immediate feedback on crop health.
2. **Multispectral and Hyperspectral Imaging: -** The use of multispectral and hyperspectral imaging for disease detection can provide more detailed information about plant health. Future research should focus on the integration of these technologies into disease identification systems.
3. **Early Disease Detection:** - Detecting diseases at their earliest stages is crucial for effective disease management. Future studies should concentrate on improving the sensitivity and accuracy of early disease detection models.
4. **Integration with Agricultural Practices:** - Bridging the gap between disease detection and agricultural practices is essential. Researchers should explore ways to incorporate disease identification into farming processes, enabling timely intervention.
5. **Global Data Sharing: -** To improve the accuracy of disease classification models, a global database of plant diseases and associated symptoms should be established. Collaboration and data sharing among researchers worldwide can contribute to this initiative.
6. **User-Friendly Applications: -** Developing user-friendly mobile applications that allow farmers and agricultural experts to easily identify and manage plant diseases can have a significant impact on crop yields.
7. **Robustness Against Environmental Variations:** - As environmental conditions play a vital role in disease development; models should be made more robust and adaptable to varying conditions.

References

[1] S Phani P., Rajeswari N., Anuradha C., Venkata Nagaraju T., Sai Srinivas V., Uddagiri S, "A Novel Classification Approach for Grape Leaf Disease Detection Based on Different Attention Deep Learning Techniques," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 14, No. 6, 2023.

[2] Singh, G., & Yogi, K. K. (2023). Performance Evaluation of Plant Leaf Disease Detection Using Deep Learning Models. Archives of Phytopathology and Plan 209-233. https://doi.org/10.1080/03235408.2023.2183792.

[3] Ramanjot, Usha Mittal, Ankita Wadhawan, Jimmy Singla, N.Z Jhanjhi, Rania M. Ghoniem, Sayan Kumar Ray, and Abdelzahir Abdelmaboud. "Plant Disease Detection and Classification: A Systematic Literature Review." Sensors 2023, 23, 4769. https://doi.org/10.3390/s23104769.

[4] Mahum, R., Munir, H., Mughal, Z.-U.-N., Awais, M., Khan, F. S., Saqlain, M., Mahamad, S., & Tlili, I. (2022). A novel framework for potato leaf disease detection using an efficient deep learning model. Human and Ecological Risk Assessment: An International Journal, 303-326. https://doi.org/10.1080/10807039.2022.2064814

[5] Pandian, J.A.; Kanchanadevi, K.; Kumar, V.D.; Jasinska, E.; Gono, R.; Leonowicz, Z.; Jasinski, M. A Five Convolutional Layer Deep Convolutional Neural Network for Plant Leaf Disease Detection. Electronics 2022, 11, 1266.

[6] Vallabhajosyula, S.; Sistla, V.; Kolli, V.K.K. Transfer learning-based deep ensemble neural network for plant leaf disease detection. J. Plant Dis. Prot. 2022, 129, 545–558.

[7] Ashwinkumar, S.; Rajagopal, S.; Manimaran, V.; Jegajothi, B. Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks. Mater. Today Proc. 2022, 51, 480–487.

[8] Sharma, R.; Singh, A.; Jhanjhi, N.Z.; Masud, M.; Jaha, E.S.; Verma, S. Plant disease diagnosis and image classification using deep learning. CMC-Comput. Mater. Contin. 2022, 71, 2125–2140.

[9] Ghosh, S.; Singh, A.; Kavita Jhanjhi, N.Z.; Masud, M.; Aljahdali, S. SVM and KNN Based CNN Architectures for Plant Classification. CMC-Comput. Mater. Contin. 2022, 71, 4257–4274.

[10] Malathy, S.; Karthiga, R.; Swetha, K.; Preethi, G. Disease Detection in Fruits Using Image Processing. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20–22 January 2021; pp. 747–752.

[11] Wassan, S.; Xi, C.; Jhanjhi, N.Z.; Binte-Imran, L. Effect of frost on plants, leaves, and forecast of frost events using convolutional neural networks. Int. J. Distrib. Sens. Netw. 2021, 17, 15501477211053777.

[12] Bedi, P.; Gole, P. Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network. Artif. Intell. Agric. 2021, 5, 90–101.

[13] Chouhan, S.S.; Singh, U.P.; Jain, S. Automated Plant Leaf Disease Detection and Classification Using Fuzzy Based Function Network. Wirel. Pers. Commun. 2021, 121, 1757–1779.

[14] Chellapandi, B.; Vijayalakshmi, M.; Chopra, S. Comparison of Pre-Trained Models Using Transfer Learning for Detecting Plant Disease. In Proceedings of the 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 19–20 February 2021; pp. 383–387.

[15] Chen, L.; Cui, X.; Li, W. Meta-Learning for Few-Shot Plant Disease Detection. Foods 2021, 10, 2441.

[16] Sujatha, R.; Chatterjee, J.M.; Jhanjhi, N.Z.; Brohi, S.N. Performance of deep learning vs machine learning in plant leaf disease detection. Microprocess. Microsyst. 2021, 80, 103615.

[17] Jain, S.; Dharavath, R. Memetic Salp Swarm Optimization Algorithm Based Feature Selection Approach for Crop Disease Detection System. J. Ambient Intell. Humaniz. Comput. 2021, 14, 1817–1835.

[18] Jadhav, S.B.; Udupi, V.R.; Patil, S.B. Identification of plant diseases using convolutional neural networks. Int. J. Inf. Technol. 2021, 13, 2461–2470.

[19] Chowdhury, M.E.H.; Rahman, T.; Khandakar, A.; Ayari, M.A.; Khan, A.U.; Khan, M.S.; Al-Emadi, N.; Reaz, M.B.I.; Islam, M.T.; Ali, S.H.M. Automatic and Reliable Leaf Disease Detection Using Deep Learning Techniques. AgriEngineering 2021, 3, 294–312.

[20] Lijo, J. Analysis of Effectiveness of Augmentation in Plant Disease Prediction using Deep Learning. In Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 8–10 April 2021; pp. 1654–1659.

[21] Abbas, A.; Jain, S.; Gour, M.; Vankudothu, S. Tomato plant disease detection using transfer learning with C-GAN synthetic images. Comput. Electron. Agric. 2021, 187, 106279.

[22] Akshai, K.P.; Anitha, J. Plant Disease Classification Using Deep Learning. In Proceedings of the 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Coimbatore, India, 13–14 May 2021; pp. 407–411.

[23] Atila, Ü.; Uçar, M.; Akyol, K.; Uçar, E. Plant leaf disease classification using EfficientNet deep learning model. Ecol. Inform. 2021, 61, 101182.

[24] Sun, Y.; Liu, Y.; Zhou, H.; Hu, H. Plant Diseases Identification through a Discount Momentum Optimizer in Deep Learning. Appl. Sci. 2021, 11, 9468.

[25] Wang, J.; Yu, L.; Yang, J.; Dong, H. Dba_ssd: A novel end-to-end object detection algorithm applied to plant disease detection. Information 2021, 12, 474.

[26] Sharma, P., Berwal, Y. P. S., & Ghai, W. (2019). Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. Innovative Plant Science, 1, 10. https://doi.org/10.1016/j.inpa.2019.11.001.

[27] Karthik R., Hariharan M., Sundar Anand, Priyanka Mathikshara, Annie Johnson (2020). "Attention embedded residual CNN for disease detection in tomato leaves." Applied Soft Computing, 86, 105933. https://doi.org/10.1016/j.asoc.2019.105933.

[28] Dengshan Li, Rujing Wang, Chengjun Xie, Liu Liu, Jie Zhang, Rui Li, Fangyuan Wang, Man Zhou, Wancai Liu (2020). "A Recognition Method for Rice Plant Diseases and Pests Video Detection Based on Deep Convolutional Neural Network." Sensors, 20(3), 578. https://doi.org/10.3390/s20030578.

[29] Ji, M.; Zhang, L.; Wu, Q. Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks. Inf. Process. Agric. 2020, 7, 418–426.

[30] Nithish, E.K.; Kaushik, M.; Prakash, P.; Ajay, R.; Veni, S. Tomato Leaf Disease Detection Using Convolutional Neural Network with Data Augmentation. In Proceedings of the 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 10–12 June 2020; pp. 1125–1132.

[31] Chen, J.; Chen, J.; Zhang, D.; Sun, Y.; Nanehkaran, Y.A. Using deep transfer learning for image-based plant disease identification. Comput. Electron. Agric. 2020, 173, 105393.

[32] Shrestha, G.; Deepsikha; Das, M.; Dey, N. Plant Disease Detection Using CNN. In Proceedings of the 2020 IEEE Applied Signal Processing Conference (ASPCON), Kolkata, India, 7–9 October 2020; pp. 109–113.

[33] Pham, T.N.; Van Tran, L.; Dao, S.V.T. Early Disease Classification of Mango Leaves Using Feed-Forward Neural Network and Hybrid Metaheuristic Feature Selection. IEEE Access 2020, 8, 189960–189973.

[34] Marzougui, F.; Elleuch, M.; Kherallah, M. A Deep CNN Approach for Plant Disease Detection. In Proceedings of the 2020 21st International Arab Conference on Information Technology (ACIT), Giza, Egypt, 28–30 November 2020; pp. 1–6.

[35] Selvam, L.; Kavitha, P. Classification of lady’s finger plant leaf using deep learning. J. Ambient Intell. Humaniz. Comput. 2020.

[36] M, Turkoglu.; Hanbay, D. Plant disease and pest detection using deep learning-based features. Turk. J. Electr. Eng. Comput. Sci. 2019, 27, 1636–1651.

[37] Too, E.C.; Yujian, L.; Njuki, S.; Yingchun, L. A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric. 2019, 161, 272–279.

[38] Picon, A.; Alvarez-Gila, A.; Seitz, M.; Ortiz-Barredo, A.; Echazarra, J.; Johannes, A. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. Comput. Electron. Agric. 2019, 161, 280–290.

[39] Arsenovic, M.; Karanovic, M.; Sladojevic, S.; Anderla, A.; Stefanovic, D. Solving Current Limitations of Deep Learning Based Approaches for Plant Disease Detection. Symmetry 2019, 11, 939.

[40] X., Zhang; Han, L.; Dong, Y.; Shi, Y.; Huang, W.; Han, L.; González-Moreno, P.; Ma, H.; Ye, H.; Sobeih, T. A Deep Learning-Based Approach for Automated Yellow Rust Disease Detection from High-Resolution Hyperspectral UAV Images. Remote Sens. 2019, 11, 1554.

[41] Kumari, C.U.; Prasad, S.J.; Mounika, G. Leaf Disease Detection: Feature Extraction with K-Means Clustering and Classification with ANN. In Proceedings of the 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 27–29 March 2019; pp. 1095–1098.

[42] Al Haque, A.S.M.F.; Hafiz, R.; Hakim, M.A.; Islam, G.M.R. A Computer Vision System for Guava Disease Detection and Recommend Curative Solution Using Deep Learning Approach. In Proceedings of the 2019 22nd International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 18–20 December 2019; pp. 1–6.

[43] Francis, M.; Deisy, C. Disease Detection and Classification in Agricultural Plants Using Convolutional Neural Networks—A Visual Understanding. In Proceedings of the 2019 6th international Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 7–8 March 2019; pp. 1063–1068.

[44] Sahithya, V.; Saivihari, B.; Vamsi, V.K.; Reddy, P.S.; Balamurugan, K. GUI Based Detection of Unhealthy Leaves Using Image Processing Techniques. In Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 4–6 April 2019; pp. 0818–0822.

[45] Howlader, M.R.; Habiba, U.; Faisal, R.H.; Rahman, M.M. Automatic Recognition of Guava Leaf Diseases Using Deep Convolution Neural Network. In Proceedings of the 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox’sBazar, Bangladesh, 7–9 February 2019; pp. 1–5.