International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

PLANT LEAF DISEASE DETECTION

KUSH VIJAY SOMAIYA
B.Tech Electronics and Computer Science Engineering
Shah And Anchor Kutchhi Engineering college
kushsomaiya0@gmail.com

HEET SHAH
B.Tech Electronics and Computer Science Engineering
Shah And Anchor Kutchhi Engineering college
heets1502@gmail.com

MIHIKA METHA
B.Tech Electronics and Computer Science Engineering
Shah And Anchor Kutchhi Engineering college
Mehta.mihika76@gmail.com

RISHI MEVADA
B.Tech Electronics and Computer Science Engineering
Shah And Anchor Kutchhi Engineering college
rishimevada07@gmail.com

SALABHA JACOB
Assistant Professor
Shah And Anchor Kutchhi Engineering college
salabha.jacob@sakec.ac.in

***ABSTRACT** - Agriculture remains a fundamental pillar of many national economies, making the protection of crops from disease a top priority. Pathogens such as bacteria, fungi, and viruses can significantly reduce crop productivity, underscoring the need for timely and accurate disease detection. Recent innovations in computer vision and artificial intelligence have introduced powerful tools for recognizing plant diseases through image analysis, particularly using leaf imagery. This paper investigates the application of machine learning, deep learning, and few-shot learning models in automating disease identification to assist farmers in making informed, prompt decisions. By examining the use of advanced models—including convolutional networks and vision transformers—alongside imaging technologies like hyperspectral cameras, this study highlights both the technological advancements and their potential impact in the field. Furthermore, it touches on molecular-level diagnostic techniques aimed at minimizing the threat of pathogens. The review offers a thorough overview of current progress and identifies key opportunities for future research, with the goal of translating laboratory breakthroughs into practical solutions for sustainable agriculture.*

INDEX TERMS : Plant disease, deep learning, machine learning, shot learning, computer vision, folding networks (CNNS), vision trans, hyperspectral imaging, molecular diagnostics, sustainable agriculture detection.**.**

I. INTRODUCTION

For centuries, agriculture has been fundamental to the growth of civilizations and remains a vital industry that supports the livelihood of millions across the globe. Today, nearly one billion people work in agriculture, contributing significantly to national economies and ensuring food availability. Yet, plant diseases pose a serious challenge, leading to annual losses of more than 10% in crop yields. This issue adds to the growing concern of global hunger, which currently affects around 680 million people and continues to rise. [1]

Plant disease emergence is generally caused by a disease-causing organism, including fungi, viruses and bacteria. Some disease symptoms which are transmitted to plants appear on the surface of a plant, especially leaves. While many diseases will develop inside a plant or will not show signs of infection until significant damage has occurred. Traditional diagnosis methods can take place through visual inspection, or laboratory based molecular methods such as ELISA and PCR. While traditional methods of disease diagnosis are highly accurate, they are often limited in scope, expensive, complicated and still require an expert to conduct the work. [2] All of these things make them an inefficient choice for occasion use in fields larger than backyard gardens or for farmers with limited resources. [3]

As an alternative to these traditional methods, AI has emerged as a potential solution to self automate plant disease identification. Of the many types of AI methodologies, deep learning and computer vision have shown early success in identifying patterns of disease in leaf images. [4] These methods are fast, reliable and non-invasive and may facilitate farmers recognizing potential diseases at an earlier stage and ensure the expenditures that follow as a result of undetected or misdiagnosed damage from a plant disease are mitigated. [5]

Deep learning models are generally very effective for image-based classification problems, particularly in the case of convolutional neural networks (CNNs) as well as more recent

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

models such as vision transformers and few-shot learning (FSL) models. [6] These models reduce the need for manual feature extraction, as they can learn features from raw data, and they reduce reliance on extensive amounts of labeled data through synthetic data generation and data augmentation that can enable models to learn how to better generalize in real-world situations. [7]

However, significant challenges remain when deploying new technologies across a wide diversity of landscapes. For example, the models that are created in laboratory environments tend to perform poorly in the real world because of different lighting, background noise, and variability in the environment. [8] Deep learning models, in general, require large training set size and require high computational power, which are barriers for low-resource settings. Few-shot learning provides a relatively low-cost alternative by enabling the model to learn from a limited number of labeled samples, enabling a more scalable solution to disease detection. [9]

This paper reviews the applications of deep learning and computer vision in the detection of plant diseases. [7], [8] As it compares the most recent and advanced models of AI, and techniques and applications in deep learning and computer vision, it also highlights recent advancements in FSL that are helping mitigate logistical issues. This paper seeks to document the current use of AI in agriculture to facilitate the cultivation of digital, efficient, and inclusive systems for the management of diseased crops. [9], [10]

Ultimately, this study will provide an understanding of the advancements in this area of research and elaborate on the ability of AI to improve food security in an inclusive and impactful way for farmers and sustainable agriculture. [11], [12]

II. PHYTOPATHOLOGY

Phytopathology is the branch of science dedicated to studying plant diseases, their underlying causes, how they spread, and methods for managing their impact on crops. Derived from the Greek words for plant (*phyto*), disease (*patho*), and study (*logos*), this field focuses on protecting plant health throughout their life span. [13]

The discipline aims to determine whether plant diseases are triggered by living organisms such as fungi, bacteria, or viruses (biotic causes), or by environmental stressors like poor soil or extreme weather (abiotic causes). It also examines how diseases progress (pathogenesis), how they interact with host plants (epidemiology), and how to effectively manage or minimize crop loss. Fig.1

Phytopathology combines insights from a wide range of sciences including microbiology, botany, virology, meteorology, climatology, genetic engineering, and molecular biology. By merging these fields, it helps in understanding and developing strategies to combat plant diseases, ultimately ensuring better crop yields and contributing to global food security. Fig.2

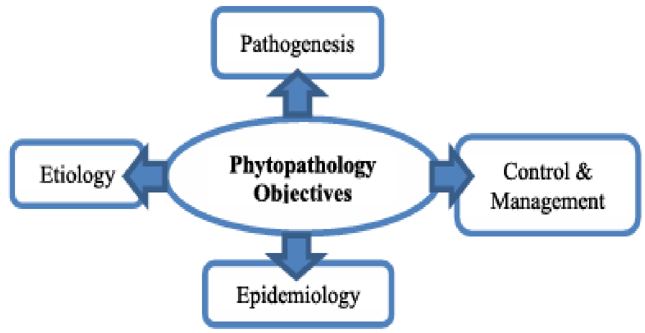


FIGURE 1: PHYTOPATHOLOGY OBJECTIVES.

FIGURE 2: SUBDOMAINS OF PHYTOPATHOLOGY [13]

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

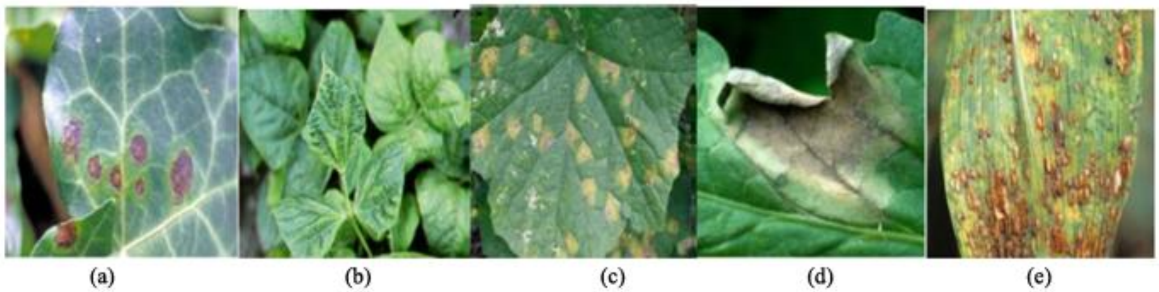


FIGURE 4: (A) BACTERIAL BLEMISH (B) VIRAL MOSAIC (C) LATE BLIGHT (D) EARLY BLIGHT (E) RUST

III. PLANT DISEASE/TYPES AND SYMPTOMS

When a plant's natural functions or development are disrupted, it often indicates the presence of disease. These abnormalities can be triggered by two main types of factors: biotic (living organisms) or abiotic (non-living environmental conditions), as shown in Figure 3 [13]. While abiotic causes—such as drought, nutrient deficiency, or pollution—are usually not contagious and can often be avoided, this paper focuses primarily on biotic diseases, which are caused by infectious agents.

1) Bacterial Infections

Bacterial diseases in plants usually begin with small, wet-looking spots that eventually turn into dry, discolored patches (see Figure 4a). These may appear as dark or brown spots, or as black blemishes surrounded by a yellowish halo. In dry weather, they can develop a speckled look. A common example is bacterial wilt in brinjal (eggplant), where the entire plant may suddenly collapse due to infection. [14]

2) Viral Infections

Viral plant diseases are particularly challenging to diagnose because they often show no clear symptoms or mimic signs of chemical damage or nutrient loss. Common carriers of these viruses include beetles, aphids, whiteflies, and leafhoppers [14]. One well-known example is the mosaic virus, which causes yellow or green streaks on leaves (see Figure 4b). Because the symptoms are so subtle, viral diseases often go unnoticed until serious damage has occurred.

3) Fungal Infections

Mushroom pathogens affect a wide range of plant parts, including stems, leaves, seeds and roots. Examples are stem rack, silk, wilt, ergot and blackpoint. A critical disease, delayed failure caused by plant spheres, usually begins with grey green spots in the lower lobe (Fig. 4C). These stains can be soaked in water and darkened over time, with white mushroom spots that can eventually grow on the surface, especially in wet weather conditions. [15]

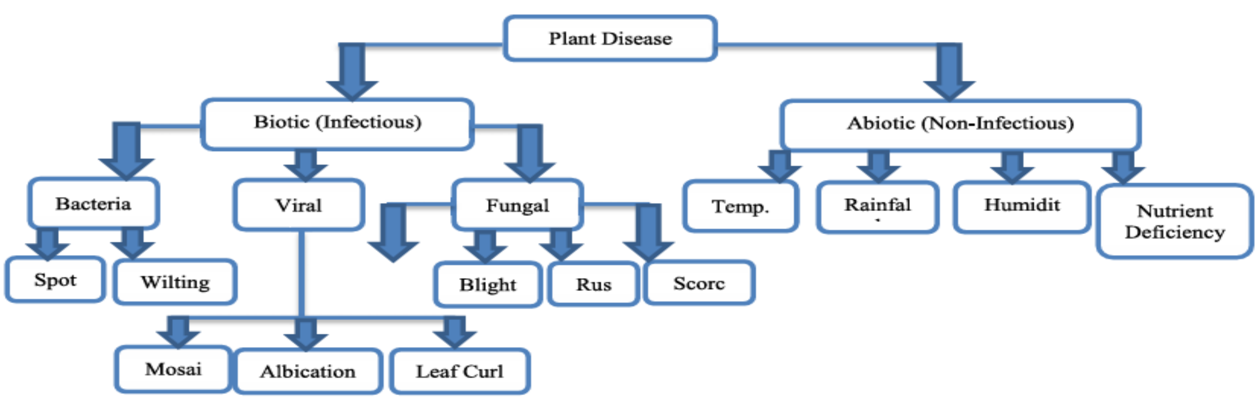


FIGURE 3 : CLASSIFICATION OF PLANT DISEASE IN DISTINCT CATEGORIES [13]

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Another example, early blight, caused by the *Alternaria* fungus, shows up as small brown spots with circular ring patterns, resembling a bull's eye (Figure 4d). Rust fungus, on the other hand, appears on mature leaves, starting as yellow or greenish patches that later turn dark brown or black (Figure 4e).

Limitations of Symptom-Based Plant Disease Diagnosis

The symptoms mentioned above are just a small sample of those described for plant diseases. Often the same or very similar symptoms can be caused by both infectious (biotic) and non-infectious (abiotic) pathogens that, many times, it becomes difficult to distinguish them from one another. Because of this overlap, identifying the particular disease or pathogen who such symptoms belong to, is often inaccurate when only utilizing symptoms. Table 1 lists some identifiable plant diseases that may assist for identifying these patterns.

TABLE 1 : DISTINCT PLANTS THEIR DISEASE AND RESPONSIBLE PATHOGEN [13]

Plant	Disease	Pathogen	Symptoms
Apple	Scab	Pomi Spilocaca	Brown-Gray on leaf
	Rot	Malorum Sphaeropsis	Dark Brown on leaf
	Rust	Sporangium	Yellow pale on leaf
Cherry	Mildew	Clandestina	Gray powder on leaf
Corn	Gray Spot	Cercospora	Rectangle lesions
	Rust	Sorghi puccinia	Red pustules on leaf
	Light blight	Tutcica setosphaeria	Elliptical lesions
Grape	Rot	Bidwellii guignardia	Red borders on leaf
	Measles	Aleophilum	Necrotic stripping
	Isariopsis blight	Angulata brachypus	Coalesce lesions
Peach	Spot	Arboricola Xanthomonas	Clustered lesions
Potato	Early blight	Solani Alternaria	Brown lesion
	Late blight	Infestans phytophthora	Dark greeb spot
Tomato	Septoria spot	Lycopersici	Foliage
	Mosaic	Mosaic virus	Mottle green leaf
Orange	Green Citrus	Bacteria Motile	Precipitate Demolition
Strawberry	Scorch Fungus	Diplocarpon	Brown edges
Squash	Mildew	Xanthii podosphaers	White powder

Diagnosis of plant diseases accurately can be very challenging, especially when the symptoms are either non-specific or vague. Visual assessments alone may not be sufficient to determine the underlying cause and causal agent.

To overcome these challenges, artificial intelligence (AI) techniques have gained traction in plant disease detection. [16], [17], [18].These systems largely depend on two fundamental processes: extracting relevant features from input data and classifying diseases based on those features.

By identifying and analyzing key characteristics—such as color shifts, shape anomalies, or texture changes—AI models can differentiate between various plant diseases with increased precision.

The features selected for analysis serve as indicators of specific diseases. Table 2 presents several common plant diseases along with their corresponding symptoms, offering a reference point for researchers. Leveraging these features can significantly enhance classification accuracy, making the overall disease detection process more reliable and effective.

IV. SELECTION PROCESS

1) Creating Search Queries

These keywords were then combined to form strategic search queries designed to retrieve a broad yet relevant range of studies. For instance, search strings such as "plant disease detection AND deep learning" or "plant disease detection OR computer vision" were used to capture publications across multiple intersecting fields. [19]

2) Searching Databases

Comprehensive searches were performed across a number of respected academic databases. These included PubMed, Semantic Scholar, Scopus, Google Scholar, IEEE Explore, Science Direct, and Web of Science. These platforms were chosen for their relevance to domains like computer science, agricultural engineering, and AI research, and for their extensive collections of peer-reviewed articles. [19]

3) Preliminary Screening

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

In the initial screening phase, titles and abstracts were reviewed to quickly assess the relevance of each article. Studies that did not meet the predefined inclusion criteria were filtered out at this stage, helping reduce the volume of results to a more focused and manageable set for deeper analysis. [19]

4) In-Depth Assessment

The remaining studies were subjected to a detailed full-text analysis. Each paper was assessed for its research objectives, methodology, overall quality, key findings, and its specific contributions to the field. This step ensured that only studies with substantial relevance and academic rigor were retained

5) Final Inclusion

After all screening criteria were used, a final list of 278 research articles was created. This structured, systematic approach has formed a solid foundation for understanding modern development and potential future directions when using deep learning and computer vision to identify plant diseases in precision agriculture.

V. PLANT DISEASE DETECTION SYSTEM

Artificial Intelligence (AI) technologies are increasingly being used to support agricultural productivity by enabling accurate monitoring of plant health. Although numerous review studies have been conducted—some targeting specific methodologies and others focusing on particular plant diseases—there remains a lack of comprehensive reviews that combine detection, classification, and diagnostic strategies into a single framework. [20]

This review aims to bridge that gap by examining diverse methodologies employed by researchers. These include the use of machine learning (ML), deep learning (DL), few-shot learning (FSL), and soft computing techniques, integrated with image processing tools that utilize both RGB and hyperspectral imaging. Additionally, it highlights molecular-based methods developed to identify and manage the threat posed by plant pathogens. [20]

A. Machine Learning with Image-Based Analysis

The process of detecting plant diseases using machine learning is typically divided into several key steps: Image Covering, preprocessing, image segmentation, distinctive extraction and selection, and classification. Each stage plays an important role in the performance of the discrimination model, and various methods have been proposed in the literature to optimize these phases.. [21]

TABLE 2: DISTINCT DISEASE IN DIFFERENT PLANTS

Author's	Plant Name	Bacterial Disease	Viral Disease	Fungal Disease
Zhang et al. 2019 [19] Kianat et al. 2021 [20] Agarwal et al. 2021 [21]	Cucumber	Brown blemish, Angular Blemish, Target blemish	Mosaic, Yellow blemish	Black blemish, Gray mold
Shrivastava et al. 2019 [22] Chen et al. 2021 [23]	Rice	Streak, Blight	Black Dwarf Streaked	Smut False
Sun et al. 2021 [24]	Maize	Streak, Stalk	Crimson, Dwarf	Rust
Ferentinos 2018 [25] Abbas et al. 2021 [26]	Tomato	Canker	Curl leaf yellow	Late/ Early Blight

1) Image Acquisition

The initial stage in developing a machine learning system for disease detection involves gathering images of plant components such as leaves, stems, roots, or branches. The accuracy of detection largely depends on the quality of these images, which is influenced by the type of camera and environmental settings. [22]

Images captured in natural, uncontrolled environments may contain unwanted elements like shadows, cluttered backgrounds, or noise. Therefore, removing these distortions—particularly background elements and visual noise—is crucial in improving detection accuracy. In addition to standard RGB cameras, researchers also employ specialized imaging tools that capture hyperspectral, thermal, and fluorescence data for enhanced analysis.

Various plant disease datasets used in research are summarized in Table 3. The inconsistency in lighting and background complexity found in field conditions often contrasts with the ideal settings of laboratory environments, which directly affects the efficiency of the detection process. Hence, robust image acquisition plays a pivotal role in system accuracy.

2) *Image Preprocessing*

Image preprocessing is a foundational step in the machine learning workflow. It is responsible for enhancing the visual quality of images degraded by environmental factors like poor lighting or motion blur. Since many datasets are collected in real-time agricultural settings, preprocessing helps prepare the images for accurate feature extraction and minimizes the computational load. [22]

Common pre-processing steps include cropping, resizing, and contrast adjustment, as well as the removal of unwanted backgrounds. The choice of techniques varies depending on the image quality. An overview of pre-processing strategies applied by different studies can be found in Table 4.

Additionally, image augmentation techniques are employed to expand existing datasets—an important practice in training deep learning models that require large quantities of labelled data. Techniques used include random flipping, noise addition, rotation, scaling, gamma correction, zooming, shifting, and other image transformations such as brightness and contrast enhancements.

3) *Image Segmentation*

Segmentation is the method used to isolate affected areas of a plant image, enabling focused analysis of diseased regions. This helps separate healthy tissue from infected areas, simplifying the classification process and improving model performance.

The segmentation process, however, faces challenges such as unclear boundaries, lighting inconsistencies, and complex backgrounds. Approaches to segmentation can be broadly categorized into two types: traditional methods like thresholding, region-growing, and edge detection; and computational methods such as fuzzy logic, neural networks, and genetic algorithms. In most cases, computational techniques offer more robust and accurate results than conventional ones. [23]

Effective segmentation is critical for accurate feature extraction, which directly impacts classification accuracy. A variety of segmentation approaches used in recent studies are presented in Table 5.

TABLE 3 : DETAILS OF THE DATASET AVAILABLE USED BY VARIOUS RESEARCHERS

	Dataset Name	Authors
Open accessible dataset	APS image dataset	Mohanty et al. 2016 [33]
	Plant Village Image dataset	Mohanty et al. 2016 [33]
	Computers and Optics in Food Inspection (Cofi) laboratory image dataset	Arnal Barbedo et al. 2019 [34]
	Digipathos images (PDDB)	Arnal Barbedo et al. 2019 [34]
	IRRI Dataset	Bashir et al. 2019 [35]
	INIBAP leaf Dataset	Camargo & Smith 2009 [36]
Self-created dataset	-	Karadag et al. 2018 [37]
	-	Coulibaly et al. 2019 [38]
	-	Pantazi et al. 2019 [39]
	-	Fuentes et al. 2017 [40]
	-	Shrivastava and Hooda 2014 [41]
	RoCoLe	J. Parraga Alava. et al. 2019 [42] K. Tian, et al. 2019 [43]
Multiple Crop dataset	Citrus dataset	Zhang and Meng 2011 [44]
	Grapefruit Grove	Pydipati et al. 2006 [45]
	-	Masazhar and Kamal 2018 [46]
University Agriculture	-	Deshapande et al. 2019 [47] Pujari et al. 2016 [48]
	-	Abed and Esmaeel 2018 [49]
Dataset using spectral devices	-	Azadbakht et al 2019 [50]
	-	Rothe and Kshirsagar 2015 [51]
Hyperspectral imaging	-	Abdulridha et al 2019 [52]
	-	Zhang et al 2018 [53]
Charge Couple device camera	-	Huang 2007 [54]
	-	Yao et al 2009 [55]

TABLE 4 : DETAILS OF PRE-PROCESSING TECHNIQUE USED BY VARIOUS RESEARCHERS

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Techniques Used		Author's
Color Space Conversion	Enhancement Filtering, Background reduction	Kaur et al. 2018 [56] Das et al. 2020 [57]
	RGB, HSV, HSI, YIQ, L*a*b, grayscale	Khot et al. 2016 [58], Dhingra et al. 2017 [59], Cruz et al. 2018 [60]
	L a* b*, HSV	Rothe and Kshirsagar 2015 [37], Vidyaraj and Priya 2016 [61], Kaur et al. 2018 [37], Pantazi et al. 2019 [39]
	YCbCr, CIE	Kai et al. 2011 [62], Chaudhary et al. 2012 [63], Joshi and Jadhav 2017 [64]
	H, I3a, I3b	Camargo and Smith 2009a [36]
	CIELuv	Ganeshan et al. 2017 [65]
	La*b*, Luv, YCbCr	Meunkaewjinda et al. 2008 [66]
Image Enhancement Techniques	Denoising using Mean and median filtering	Yao et al. 2009 [41], Rao & Kulkarni 2020 [67]
	Sharpening using Gaussian and Laplacian filtering	Asfarian et al. 2013 [68], Abed and Esmaeel 2018 [49]
	Illumination variation using histogram equalization	Dange and Sayyad 2015 [69], Khirade and Patil 2015 [70], Malika and Vasanthi 2017 [71]
	Augmentation	Sladojevic et al. 2016 [72], Goncharov et al. 2019 [73]

TABLE 5: DETAILS OF SEGMENTATION TECHNIQUE USED BY VARIOUS RESEARCHERS.

Techniques Used		Author's
Edge-based segmentation	Canny edge detection, Sobel Operator, Prewitt Operator	Pydipati et al. 2006 [45], Anthonys and Wickramarachchi 2009 [75], Bankar et al. 2014 [76], Shinde et al. 2015 [77]
Thresholding Techniques	Otsu Thresholding, Adaptative Thresholding, Entropy Thresholding	Khirade and Patil 2015 [70], Pujari et al. 2016 [48], Cruz et al. 2018 [60], Das et al. 2020 [57]
Region Growing	Local Threshold	Pang et al. 2011 [78], Singh et al. 2015a [79]
Clustering	k-means	Rastogi et al. 2015 [80], Jadhav and Patil 2016 [81], Zhang et al. 2017b [82], Kaur et al. 2018b [56], Bashir et al. 2019 [35] Harakamanavar et al. 2022 [83] Zhang et al. 2018 [84]
	Fuzzy c-means	Jagtap and Hambarde 2014 [85], Bai et al. 2017 [86] Sahu & Pandey 2023 [87]
Grab Cut	-	Pantazi et al. 2019 [39]
Genetic Algorithm	-	Singh et al. 2015b [88], Singh and Misra 2017 [89]

4) Feature Extraction

Feature extraction is a critical step in computer vision and machine learning systems, particularly when differentiating between different regions in an image prior to classification. The features that are extracted are also a critical step when identifying and analyzing objects and determining the correct class or category.

In the plant disease detection literature, features like shape, color, and texture are frequently used to recognize the symptoms or pattern. The overall efficacy of disease classification systems is predicated on efficient and effective feature extraction methods. Review and analysis of the literature found that the shape, texture, and color traits present in the visibly diseased areas of the leaf are the most critical and should be extracted to ensure reliable outcomes [23].

That said, ultimately identifying an appropriate set of features is not always easy. For example, many diseases in plants exhibit similar visual characteristics which makes it difficult to distinguish between one disease and another and thus the process of determining the most meaningful and relevant

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

features is an important task in developing disease diagnostic systems. [22]

In addition to extraction, feature selection is also useful in improving model performance. When selecting features for machine learning applications, the aim is to use the features that are most useful and that help avoid overfitting but also can limit the computational overhead of the system. To select the features that are relevant from a larger set of extracted features using advanced algorithms like Principal Component Analysis (PCA) and genetic algorithms or particle swarm optimization techniques.

In Table 6 we compiled the various feature extraction methods used by researchers. These methods have different benefits depending on the crop and the modality of the image used. The last 12 years of research shows the various trends in the use of different feature extraction methods on a variety of crops, demonstrated in Figure 5.

Although research into plant disease detection has come a long way, there are many unexplored feature extraction methods that need to be investigated. When selecting crops for the purpose of study, researchers often have to consider the datasets that they have and the expert validation that is available. It is important to keep testing and developing feature extraction methods, whilst improving the accuracy and adaptability of the disease detection models. [22]

5) Classifications

Classification is one of the most important and consequential stages of any computer vision or machine learning pipeline—particularly for plant disease identification. The success of this step depends on the adherence to and quality of the previous steps, such as collecting images, preprocessing, and identifying features. During this stage, a dataset has been prepared, and a model has been trained, and it involves predicting on new images if a visual indicator of whether a plant is healthy or infected. Machine learning (ML), one aspect of artificial intelligence (AI), allows computer programs to learn independently from data and improve continually over time, without having to program the instructions for every task. Because it can learn and adapt to new scenarios, ML is an appropriate fit for complex environments, such as agriculture. [22]

ML techniques fall into three primary categories:

• Supervised learning, where the system is trained on labeled data.

• Unsupervised learning, which finds patterns in unlabeled datasets.

• Semi-supervised learning, which blends both approaches for better adaptability in scenarios with limited labeled data.

TABLE 6: DETAILS OF FEATURE EXTRACTION TECHNIQUE USED BY VARIOUS RESEARCHERS

Techniques Used		Author's
Feature Descriptor	GLCM, Wavelet Transform, Haralick feature, Gabor Transform, Local Binary Patterns	Mokhtar et al. 2015 [91],
	SURF	Bhagat & Kumar 2023 [92]
Texture Feature	GLCM Features	Bashish et al. 2011 [93], Tian et al. 2014 [94], Mainkar et al. 2015 [95], Islam et al. 2017 [96] Abed and Esmaeel 2018 [49], Sharif et al. 2018 [97], Deshapande et al. 2019 [47],
	Scale-invariant feature transform	Dandawate and Kokare 2015 [98], Mohan et al. 2016 [99], Ramesh et al. 2018 [100]
	Local Binary Pattern	Waghmare et al. 2016 [101], Singh et al. 2019 [102]
	Gabor Filter Transform	Gulhane & Gurjar 2011 [103] Prasad et al. 2012 [104] Kulkarni & R.K. 2012 [105] Jolly & Raman 2016 [106] Kaur et al. 2018b [56] Rao & Kulkarni 2020 [67]
	Wavelet Transform	Sabrol & Kumar 2016a [107] Gawali et al. 2017 [108]
	Histogram of Oriented Gradient	Dalal et al. 2005 [109] Bai et al. 2009 [110] Sannakki et al. 2013 [111] Pires et al. 2016 [112] Ramesh et al. 2018 [100] Kusumo et al. 2018 [113] Zamani et al. 2022 [114]
Color Feature	Color co-occurrence matrix	Pydipati et al. 2006 [45] Kai et al. 2011 [62] Revathi and Hemalatha 2014b [115] Ramakrishnan & Sahaya 2015 [116] Chouhan et al. 2019 [90]
	Color histogram	Caglayan et al. 2013 [117] Ramesh et al. 2018 [118]
Shape Feature	-	Anthonys & Wickramarachchi 2009 [75] Camargo & Smith 2009b [119] Wang et al. 2012 [120] Phadikar et al.2013 [121] Joshi & Jadhav 2017 [122] Sengar et al. 2018 [123] Sahu & Pandey 2023 [87]

As outlined in Figure 6, a variety of classification models have been utilized in plant disease detection. Among the most common are:

- Support Vector Machines (SVM)

- Artificial Neural Networks (ANN)
- k-Nearest Neighbors (k-NN)

In some cases, researchers also incorporate additional logic-based or vegetation index approaches to refine predictions. ANN models, in particular, are implemented in numerous forms, such as:

- Feedforward Neural Networks
- Multilayer Perceptrons
- Backpropagation Networks
- Probabilistic Neural Networks
- Self-Organizing Maps

Each of these models brings its own strengths and is chosen based on the nature of the crop, disease complexity, and available data. A comparative overview of the classification techniques and their effectiveness across different studies is detailed in Table 7, offering insights

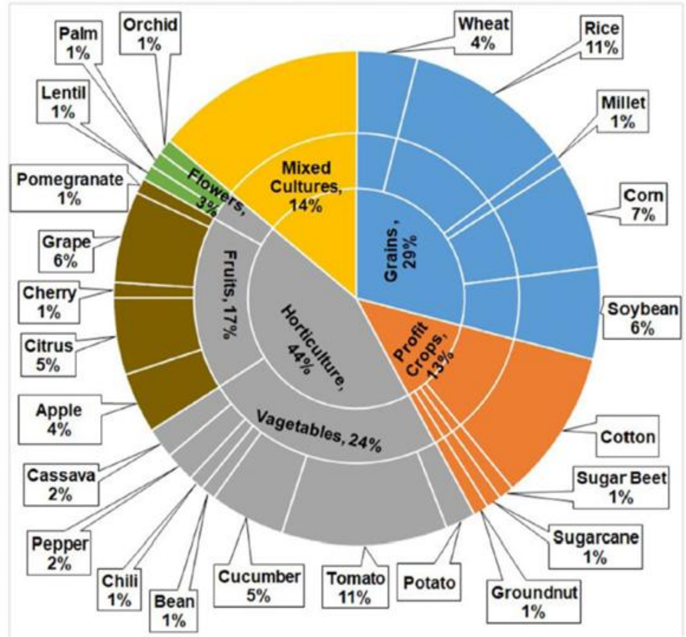


FIGURE 4: DIFFERENT CROP EXTRACTIONS

TABLE 7 : DETAILS OF CLASSIFICATION TECHNIQUE USED BY VARIOUS RESEARCHERS.

	Authors	Dataset	Feature Extraction	Classifier	Accuracy
Rice	Joshi & Jadhav 2017 [122]	Agriculture Research (115)	Color & Shape	MD & k-NN	88.15%
	Mohan et al. 2016 [99]	60	Haar & SIFT	AdaBoost, k-NN & SVM	93.33%
	Zhang et al. 2018a [53]	-	Color	Vegetation Index	63.00%
	Bashir et al. 2019 [35]	APS (440)	SIFT	SVM	94.16%
	Shrivastava & Pradhan 2021 [125]	Real Field Images	–	SVM	94.65%
	Rath & Meher 2019 [126]	Real Field Images	–	Radial Basis NN	95.00%
Wheat & Corn	Luo et al. 2015 [127]	Chinese Academy (744)	Histogram	Histogram	94.44%
	Azadbakht et al. 2019 [50]	Hyperspectral data	Index based	Regression	95.00%
	Kusumo et al. 2019 [113]	Plant Village (3823)	SIFT, SURF	SVM, DT, RF, Naïve Bayes	87.00%
	Deshapande et al. 2019 [47]	Agriculture University Dharwad (200)	First Order histogram & GLCM	k-NN, SVM	88.00%
Soyabean	Gharge & Singh 2016 [128]	IPM Database (300)	GLCM	EBPNN	93.30%
	Pires et al. 2016 [112]	Federal (1200)	SIFT, SURF, HOG	SVM	96.25%
	Kaur et al. 2018b [56]	Plant Village (4775)	Color, Texture, Shape	SVM	84.00%
Millet	Caulibaly et al. 2019 [129]	Self (124)	Transfer Learning	VGG16	89.00%
Cotton	Rothe & Kshirsagar 2015 [51]	Self A460 Camera	Hus moments	EBPNN	85.52%
	Sivasangari & Indira 2015 [130]	Self	Color, Shape	SVM	99.30%
Sugar beet	Hallau et al. 2017 [131]	Self (1400)	Texture	SVM	82.00%
Groundnut	Ramkrishnan & Sahaya 2015 [116]	Self	CCM features	EBPNN	97.41%
Cane	Pujari et al. 2016 [48]	Self (9912)	RGB Color	SVM & EBPNN	92.00%
Mix	Sladojevic et al. 2016 [72]	Internet (33469)	-	CNN	95.80%
	Ferentinos 2018 [25]	Plant Village & Self (87848)	Transfer Learning	Alexnet, VGG	99.53%
	Arnal Barbedo 2019 [34]	Self (1575)	Transfer Learning	GoogLeNet	100.00%
	Pantazi et al. 2019 [39]	-	LBP	SVM	95.00%
	Rao & Kulkarni 2020 [67]	Plant Village	Gabor, Curvelet	Fuzzy Logic	90.00%
	Ahmed & Yadav 2022 [132]	Self-Created	GLCM	Random Forest	-

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

TABLE 8 : DETAILS OF CLASSIFICATION TECHNIQUE USED BY VARIOUS RESEARCHERS.

	Sahu & Pandey 2023 [87]	Plant Vilaage	Fuzzy c means	HRF SVM	-
Apple	Samajpati & Degadwala 2016 [133]	Self (80)	Color, Histogram, LBP, Gabor	Random forest	95.00%
	Jolly & Raman 2016 [106]	Self (320)	Haarlick & LBP	SVM	96.00%
Citrus	Sharif et al. 2018 [97]	Image gallery dataset (1000)	Color, Texture, Geometrical	Multiclass SVM	95.80%
Cheery	Sengar et al. 2018 [123]	Plant Village	Lesion Area	-	99.00%
Grape	Waghmare et al. 2016 [101]	Self (450)	HSV, LBP	Multiclass SVM	89.30%
	Cruz et al. 2018 [60]	Self (272)	-	AlexNet, Inception-v3	98.00%
	Goncharov et al. 2019 [73]	Plant Village (2986)	-	Siamese	92.00%
	Javidam et al. 2023 [134]	Self	GLCM	SVM	98.97%
	Shantkumari et al. 2023 [135]	Plant Village	-	Improved k-NN	-
Pomegranate	Khot et al. 2016 [58]	-	HSI	Minimum distance classifier	-
Palm Oil	Masazhar & Kamal 2018 [46]	-	GLCM	Multiclass SVM	95.00%
Lentil	Singh et al 2019 [102]	300	LBP	Visual Examination	-
Potato	Islam et al. 2017 [96]	Plant Village (300)	Color, Texture	Multiclass SVM	95.00%
	Patil et al. 2017 [136]	Plant Village (8920)	Texture	SVM, ANN, RF	92.00%
	Verma et al 2020 [137]	Plant Village	-	Capsule Network	91.83%
Bean	Abed & Esmaeel 2018 [49]	100	GLCM	SVM	100.00%
Cucumber	Zhang et al. 2017b [82]	Self (300)	PHOG	SVM	91.48%
	Krishna Kumar & Narayan 2019 [138]	Real Field Images	-	k-means SVM	86.00%
Tomato	Raza et al. 2015 [139]	Self (71)	Pixel Value	SVM	90.00%
	Sabrol & Kumar 2016a [107]	Self (180)	Color moments, Histogram	ANFIS, FF-BPNN	87.20%
	Fuentes et al. 2017 [40]	Self (5000)	DWT, Haar Wavelet	ResNet-50, VGG16	83.06%
	Ashqar & AbuNaser 2018 [140]	Plant Village (9000)	Transfer Learning	CNN	99.84%
	Karadag et al. (2018) [37]	ARI (80)	Wavelet	ANN, NB, k-NN	84.00%
	Das et al. 2020 [57]	Self-Created	Mean, Entropy	SVM, k-NN	87.60%

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

TABLE 9 : DETAILS OF CLASSIFICATION TECHNIQUE USED BY VARIOUS RESEARCHERS

Panigrahi et al. 2020 [141]	Self-Created	-	RF	79.23%
Sujatha et al. 2021 [142]	Self-Created	-	SVM	87.00%
Zamain et al. 2022 [114]	Plant Village	PCA	SVM	-
Harakamanavar et al. 2022 [83]	-	k- means	SVM	88.00%
Rahman et al. 2023 [143]	Self-Created	GLCM	SVM	88.00%
Bhatia et al. 2020 [144]	Mildew dataset	-	Extreme Learning	89.19%
Bhatia et al. 2021 [145]	Mildew dataset	-	SVM Logistic Regression	92.73%

VI. RESULTS AND ANALYSIS

The first stage of this study depended on a small dataset of 1,532 RGB images, divided into three classes. This allowed for some early experimentation, but a small quantity of data limited the model's ability to generalize. The model achieved around 82% accuracy, but this low accuracy was mainly due to the lack of variety in features and representation so the model was learning from limited reasonable cases for predicting new items. Hence, the learning capability suffered, and the model experienced frequent errors in classification.

```
Epoch 1/5
42/42 ━━━━━━━━━━━━━━━━━━━━ 0s 3s/step - accuracy: 0.3930 - loss: 4.2992
c:\Users\Snehal Shah\AppData\Local\Programs\Python\Python312\Lib\site-packages\keras\src\trainers\data_adapters\py_da
  self._warn_if_super_not_called()
42/42 ━━━━━━━━━━━━━━━━━━━━ 118s 3s/step - accuracy: 0.3959 - loss: 4.2456 - val_accuracy: 0.6833 - val_loss: 0.7086
Epoch 2/5
42/42 ━━━━━━━━━━━━━━━━━━━━ 103s 2s/step - accuracy: 0.6736 - loss: 0.7109 - val_accuracy: 0.6167 - val_loss: 0.7342
Epoch 3/5
42/42 ━━━━━━━━━━━━━━━━━━━━ 101s 2s/step - accuracy: 0.7575 - loss: 0.5644 - val_accuracy: 0.8167 - val_loss: 0.5145
Epoch 4/5
42/42 ━━━━━━━━━━━━━━━━━━━━ 99s 2s/step - accuracy: 0.7973 - loss: 0.4980 - val_accuracy: 0.7500 - val_loss: 0.5388
Epoch 5/5
42/42 ━━━━━━━━━━━━━━━━━━━━ 100s 2s/step - accuracy: 0.8127 - loss: 0.4346 - val_accuracy: 0.8000 - val_loss: 0.4716
```

FIGURE 5: EPOCH OVERVIEW WITHOUT PARAMETER CHANGE

To resolve this, a much larger dataset of about 87,000 RGB images was used. This dataset was divided into training (80%) and validation (20%). This larger dataset had 38 classes, which gave the model a larger and more represented dataset to learn from. As a result, accuracy improved significantly to 98% due to the increased variety and size of training data allowed the model to understand the various patterns of each class better and make accurate predictions.

1) Image Preprocessing Strategy

Image preprocessing played a significant role in the performance of the model. Initially, the images were resized to 225×225 pixels, but this preprocessing step lost important detail and did not allow the model to extract useful features.

Later, the image size was adjusted to 128×128 pixels which represented a balanced trade off between the details being lost and compute overrun. Cropping was also improved to ensure important portions of the images were retained.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Batch size was held constant at 32 for the entire experiment for the capacity of the ability of the machine used.

- Layer 4: 256 filters, 3×3 kernel
- Layer 5: 512 filters, 3×3 kernel

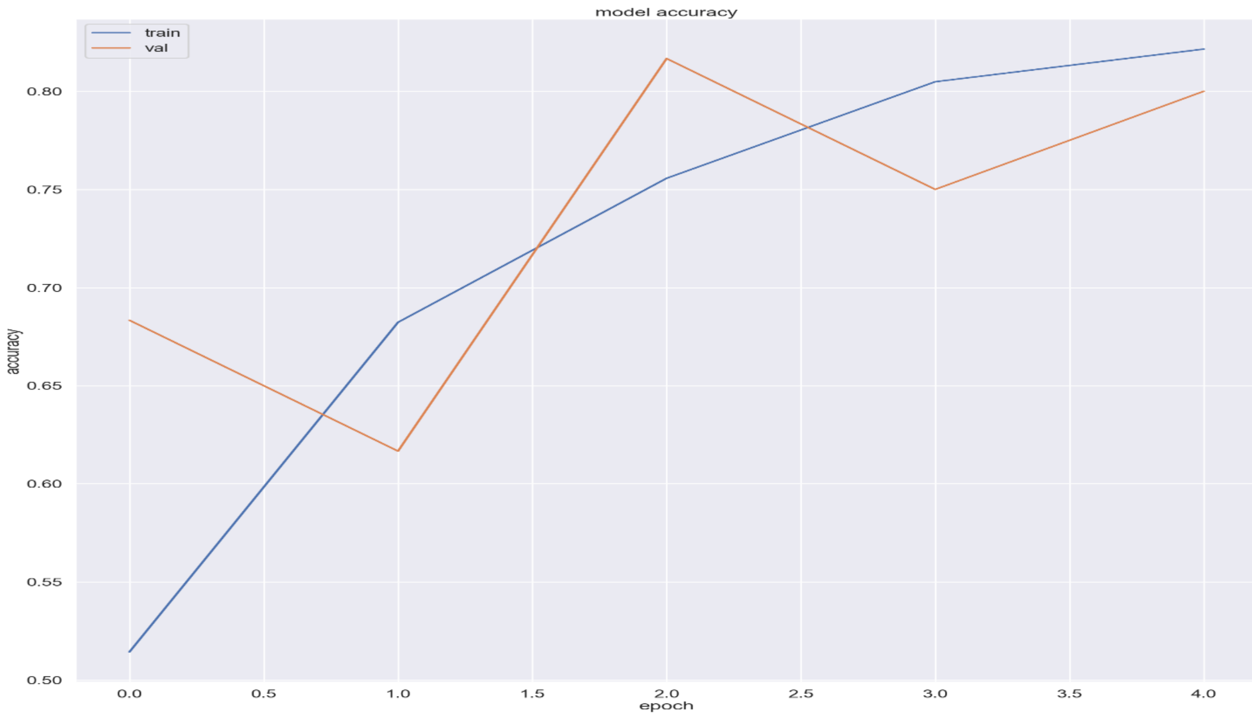


FIGURE 6: MODEL ACCURACY DURING TRAING AND VALIDATION

2) *Model Training Configuration*

The initial experiments employed five epochs of training. Due to the small dataset and low detailed input images, the five epochs produced acceptable but less-than-desired results. The limitation of epochs didn't allow for enough training depth to make more accurate future predictions.

By using a bigger and more extensive dataset and training it across 10 epochs, the final model was able to learn more complicated characteristics and perform better overall.

The present study's convolutional neural network (CNN) architecture has five convolutional layers in total, configured as follows:

- Layer 1: 32 filters, 3×3 kernel, input size of [128,128,3]
- Layer 2: 64 filters, 3×3 kernel
- Layer 3: 128 filters, 3×3 kernel

To prevent overfitting, dropout layers were integrated. Initially set at 0.25, the dropout rate was later increased to 0.4 in the final version, which helped the model generalize better across unseen data

3) *Performance Assessment*

The new model showed improved performance across many different key metrics. The precision, recall, and F1 score values were all greater, so it had better predictive ability on all of the 38 categories. Just by glancing at the confusion matrix, you could see all the data spots were accurate around the diagonal (very few errors).

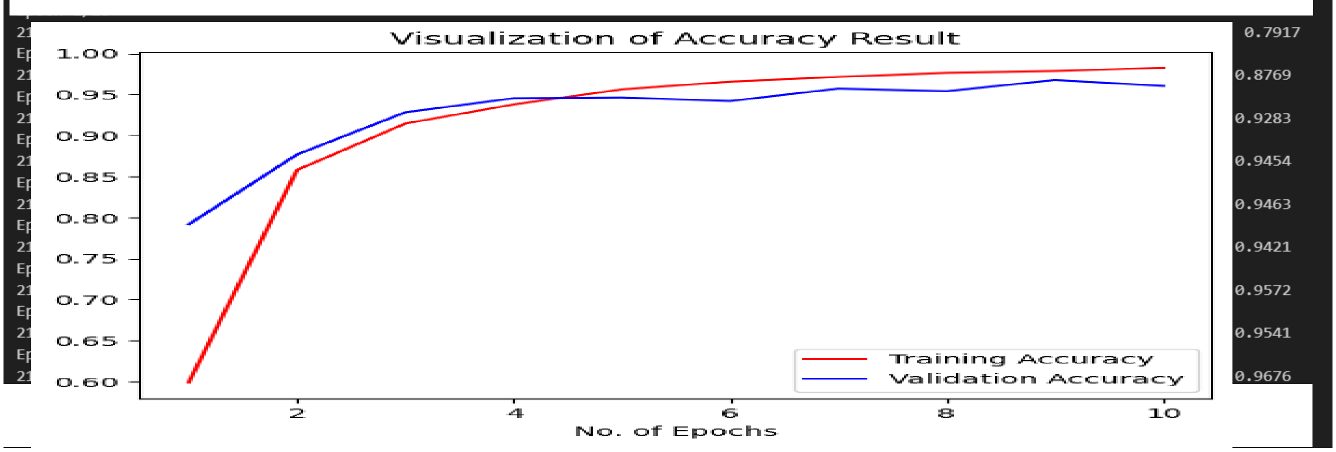
The example of accuracy increasing increases from 82% to 98% clearly illustrations the impact of (1)a larger and broader dataset, (2) improved preprocessing techniques, and (3) a deeper CNN architecture. The bottom line is this process demonstrate the advantages of continual assessment and upgrading of predictive models when developing structured deep learning models that have high accuracy.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

```
cnn.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 128, 128, 32)      896

 conv2d_1 (Conv2D)           (None, 126, 126, 32)      9248

 max_pooling2d (MaxPooling2D  (None, 63, 63, 32)       0
 )

 conv2d_2 (Conv2D)           (None, 63, 63, 64)        18496

 conv2d_3 (Conv2D)           (None, 61, 61, 64)        36928

 max_pooling2d_1 (MaxPooling  (None, 30, 30, 64)       0
 2D)

 conv2d_4 (Conv2D)           (None, 30, 30, 128)       73856

 conv2d_5 (Conv2D)           (None, 28, 28, 128)       147584

 max_pooling2d_2 (MaxPooling  (None, 14, 14, 128)      0
 2D)

...
Total params: 7,842,762
Trainable params: 7,842,762
Non-trainable params: 0
```

FIGURE 10: ACCURACY OF MODEL WHEN TRAINED

FIGURE 8: CONV2D AND MAXPOOLING LAYER WISE CONFIGURATION AND NUMBER OF PARAMETERS THAT THE MODEL WILL/LEARN UPDATE DURING TRAINING



International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

4) Website

The site was purposefully designed with a smooth, interactive, and easy-to-use way for users to navigate and find information. This means that individuals can identify plant diseases and learn about potential solutions, all in one space. The content of the site is responsive and includes a clear path to navigate through the portals, making the information accessible and pertinent to our initial project objectives.

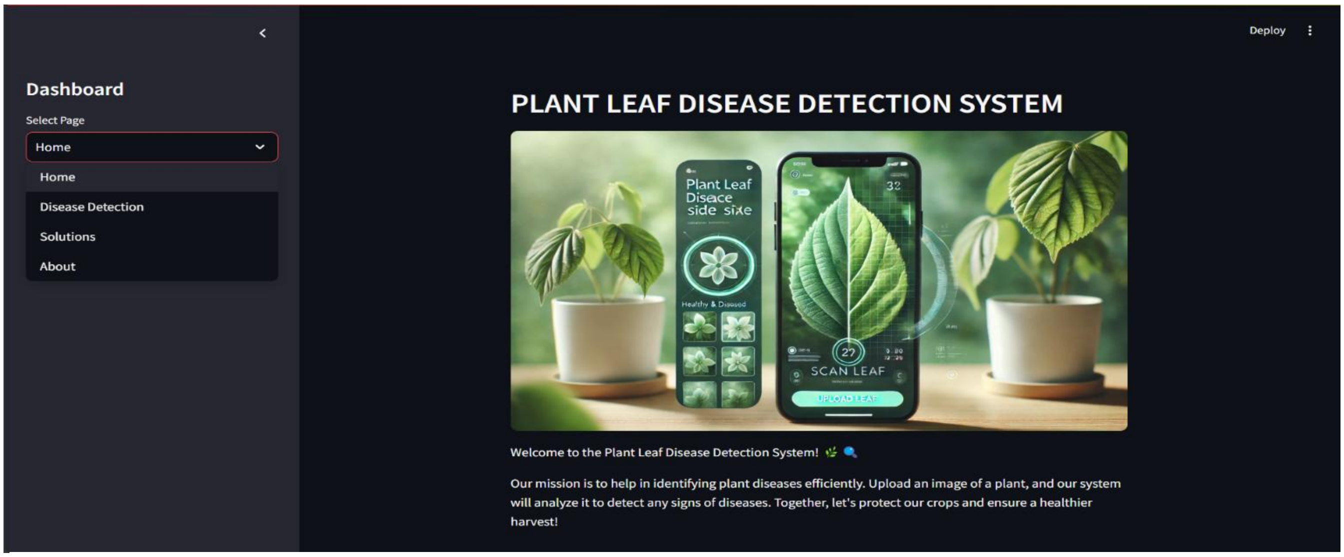


FIGURE 12: **WEBSITE HOMEPAGE**

The site combines functionality and a slick design with its python and streamlit backend. There are AI-generated images placed throughout the site for additional visual interest, and to give the site a refined, professional look.

VIII. CONCLUSION

This accounts for a clear example of how data quality, consideration in analyzing data preprocessing, and appropriate model changes can be important in the success of convolutional neural networks in image classification. At first, we trained the model on a small amount of images (1,532 RGB images), and with the short cycle of training sessions, this as expected provided 82% accuracy with no clear value due to the lack of image detail, narrow range of classes demonstrated and the number of training sessions were limited. With the change to a larger number of images to 87-000 images in total from 38 classes, and focusing on characteristic resolution of images reaching an expectable resolution of 128×128 pixels still was adequate for key visual components and belief a larger number of sessions i.e. 10 sessions was required for a change of images to input visual features for consideration. Also prior to this test, in consideration of the data each categorised images focus and theme to ensure discovery of intensive detail. The adjustments rated the performance considerably higher up to a more accurate rate from 82 % to 98%.

Within this progression from initial trial and error to polished success highlights the importance of balancing the efforts of both data-focused and architecture-focused vetted and proven improvements. Thoughtful expansion of training data, careful preprocessing methods, and sound model structure are key in developing high-performing and generalizable systems. The results of this research support the impactful use of deep learning models within practice, especially in agriculture where early and accurate detection of disease can protect crops and strengthen food systems through the world.

REFERENCES

[1] FAO; IFAD; UNICEF; WFP; WHO, "The State of Food Security and Nutrition in the World 2023," 2023.
[2] T. D. March, "State of Agriculture in India," 2023.
[3] P. Z. a. K. H. C. Janiesch, "Machine learning and deep learning," *Electron. Markets*, vol. 31, no. 3, p. 685–695, 2021.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM03128
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

[4] Z. Y. D. L. a. Z. W. K. Kc, "Impacts of background removal on convolutional neural networks for plant disease classification in-situ," *Agriculture* , vol. 11, no. 9, p. Article 827, 2021.

[5] J. L. a. X. Wang, "Plant diseases and pests detection based on deep learning: A review," *Plant Methods,* vol. 17, no. 1, pp. 1-18, 2021.

[6] M. A. a. M. A.-Z. L. C. Ngugi, "Recent advances in image processing techniques for automated leaf pest and disease recognition—A review," *Information Processing in Agriculture,* vol. 8, no. 1, p. 27–51, 2021.

[7] L. A. a. P. M. P. Chinmayi, " Survey of image processing techniques in medical image analysis: Challenges and methodologies," *Proc. Int. Conf. Soft Comput. Pattern Recognit.,* no. 10.1007/978-3-319-60618-7_45, p. 460–471, 2017.

[8] O. O. B. a. M. A. A. I. A. Adeyanju, "Machine learning methods for sign language recognition: A critical review and analysis," *Intell. Syst. Appl.,* vol. 12, p. 200056, 2021.

[9] M. Fink, "Object classification from a single example utilizing class relevance metrics," *Proc. 17th Int. Conf. Neural Inf. Process. Syst.,* p. 449–456, 2004.

[10] R. F. a. P. P. L. Fei-Fei, "One-shot learning of object categories," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 28, no. 4, p. 594–611, 2006.

[11] R. S. S. D. S. P. G. S. P. R. P. V. D. S. M. B. L. R. G. a. C. E. D. F. Martinelli, "Advanced methods of plant disease detection. A review," *Agronomy for Sustainable Development,* vol. 35, no. 1, p. 1–25, 2015.

[12] R. R. Y. Fang, "Current and prospective methods for plant disease detection," *Biosensors,* vol. 5, no. 3, p. 537–561, 2015.

[13] K. K. B. K. V. K. Vishnoi, "Plant disease detection using computational intelligence and image processing," *Journal of Plant Diseases and Protection,* vol. 128, no. 1, p. 19–53, 2020.

[14] "Evaluations of Brinjal Germplasm for Resistance to Fusarium Wilt Disease," 9 August 2023. [Online]. Available: https://www.ijsrp.org/research-paper-0717.php?rp=P676604.

[15] Y. O. D. P. P. Adhikari, "Current status of early blight resistance in tomato: An update," *Int. J. Mol. Sci.,* vol. 18, no. 10, p. 2019, 2017.

[16] K. M. S. S. K. P. H. R. Kappali, "Computer vision and machine learning in paddy diseases identification and classification: A review," *Indian J. Agricult. Res.,* vol. 10, p. 1–5, 2023.

[17] P. M. P. R. P. D. S. Joseph, "Intelligent plant disease diagnosis using convolutional neural network: A review," *Multimedia Tools Appl.,* vol. 82, no. 14, p. 21415–21481, 2022.

[18] M. T. A. Y. R. E. F. Ahmed, "Machine learning-based tea leaf disease detection: A comprehensive review," *arXiv,* 2023.

[19] N. S. C. ,. P. S. Abhishek Upadhyay, "Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture," *Springer Nature Link,* vol. 58, 2025.

[20] J. N. L. Goel, " A systematic review of recent machine learning techniques for plant disease identification and classification," *IETE Tech. Rev.,* vol. 40, no. 3, p. 423–439, 2022.

[21] A. B. A. Bhargava, "Fruits and vegetables quality evaluation using computer vision: A review," *J. King Saud Univ.-Comput. Inf. Sci.,* vol. 33, no. 3, p. 243–257, 2021.

[22] D. P. H. M. S. S. P. Mohanty, "Using deep learning for image-based plant disease detection," *Frontiers Plant Sci.,* vol. 7, p. 1419, 2016.

[23] A. M. R. E. a. C. D. S. Sankaran, "A review of advanced techniques for detecting plant diseases," *Comput. Electron. Agricult.,* vol. 72, no. 1, p. 1–13, 2010.

[24] D. P. H. a. M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostic73," 2015.

[25] A. C. a. A. P. S. S. Verma, "Prediction models for identification and diagnosis of tomato plant diseases," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI),* 2018.

