

ESS Based Citrus Leaf Classification and Disease Detection

Under the Supervision of

Dr. Ananda Kumar S

Professor

School of Computer Science and Engineering (SCOPE)

Submitted by

Team no: 05

Haarini SK – 22MID0231 – haarini.2022@vitstudent.ac.in - 6381363538

Anusha C – 22MID0277 – anusha.2022@vitstudent.ac.in – 8098677155

CSI3901 – Technical Answers for Real World Problems (TARP)

Master of Technology (Integrated)

in

Computer Science and Engineering with Specialization in Data Science



VIT[®]
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

November 2025

Citrus Leaf Classification and Disease Detection

Haarini SK, Anusha C

haarini.sk2022@vitstudent.ac.in, anusha.2022@vitstudnt.ac.in

22MID0231, 22MID0277

VIT – Vellore Institute of Technology

Abstract

Plant diseases are a major danger to agricultural production, food security, and economic stability on a global scale. Citrus crops, which are among the most economically significant fruits, are particularly susceptible to a number of bacterial and fungal diseases, such as canker, melanose, black spot, and greening. Early and accurate disease identification is crucial for sustainable farming and timely response. This paper introduces a hybrid deep learning architecture that integrates EfficientNet, Self-Organizing Maps (SOM), and Support Vector Machines (SVM) to achieve high-accuracy citrus leaf disease classification. The proposed approach combines deep feature extraction, unsupervised pattern organization, and robust classification into a single pipeline. The comprehensive preparation procedures that enhanced image quality and model generalization included denoising, contrast enhancement, edge detection, and dual segmentation (GrabCut and Otsu). Based on experimental evaluation on an enhanced dataset of 6,677 photos, the suggested model achieved an accuracy of 90.12%, outperforming six benchmark algorithms in all important measures, including precision, recall, specificity, MAE, RMSE and Cross Entropy Loss. The results validate the effectiveness of the proposed hybrid technique in providing a reliable and computationally economical diagnostic system suitable for real-time agricultural applications. This study promotes precision agriculture by assisting farmers in detecting citrus illnesses early, minimizing crop losses, and enhancing sustainable yield management.

Keywords: Citrus Leaf Disease Detection, EfficientNet, Self-Organizing Map (SOM), Support Vector Machine (SVM), Deep Learning, Image Processing, Precision Agriculture, Hybrid Model, Disease Classification, Early Diagnosis.

I. INTRODUCTION

Agriculture still plays a significant role in global economic growth in addition to generating food, guaranteeing nutritional security and sustainable lives. However, the industry is finding it more and more challenging to supply the expanding food demands of a growing global population due to crop disease prevalence, resource constraint, and climate variability. One of these problems that poses a major threat to agricultural productivity and economic stability is

plant diseases. The Food and Agriculture Organization (FAO) of the United Nations estimates that plant diseases cost the globe more than US \$220 billion annually, underscoring the necessity of effective disease detection and management strategies [1].

Citrus crops are one of the most commercially important fruit types in the world due to their high nutritional value and extensive application in the food sector. Citrus supports the livelihoods of millions of people and makes a substantial contribution to agricultural economies on every continent. Citrus output is, however, greatly impacted by a number of infectious diseases, including bacterial brown spot, melanose, citrus canker, anthracnose, Huanglongbing (HLB), and Citrus Vein Phloem Degeneration (CVPD). Significant production and quality decreases are the result of these viruses' rapid spread through environmental variables and insect vectors. Therefore, early detection and timely intervention are essential for effective crop protection and disease control [2]. Traditional plant disease diagnosis relies heavily on laboratory-based pathology analysis and expert visual inspection. Notwithstanding their efficacy, these methods have limitations with regard to subjectivity, expense, time, and accessibility.

Expert-based evaluations demand significant subject knowledge and are not scalable to large agricultural settings, whereas laboratory studies require specialized equipment and are sometimes impractical for small-scale or distant farmers [3]. Furthermore, because traditional approaches often fail to detect diseases in their early stages, when symptoms are visually modest, options for preventive intervention are limited. Recent advancements in digital imaging and precision agriculture technologies have led to the emergence of automated, data-driven disease diagnosis methods.

Early research merged traditional machine learning methods such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN) with hand created feature extraction. Although these techniques delivered good accuracy in controlled circumstances, their reliance on manually generated features hampered applicability in various backdrop, lighting, and illness situations [4].

The introduction of deep learning (DL) has revolutionized computer vision applications, enabling end-to-end feature learning directly from image data. Convolutional neural networks, or CNNs, have surpassed traditional machine learning techniques in a number of domains, including pattern recognition and hierarchical feature extraction [5]. CNNs have been useful in agriculture for detecting diseases in a range of crops, including bananas, rice, wheat, tomatoes, and apples.

Using deep learning to identify citrus illnesses still faces some challenges in spite of these advancements. First, most existing research mainly focuses on disease categorization, which is the process of assigning a picture to a predefined disease category, rather than identifying sick areas within the leaf. Although classification provides information on the presence of a disease, it does not provide spatial localization, which is crucial for assessing the disease's severity and course [6]. Bounding boxes are used in object detection frameworks to both categorize and pinpoint sick regions, thereby addressing this problem. Their application in citrus illnesses is currently limited, nevertheless.

Second, the development of robust detection models requires large, high-quality annotated datasets. The existing datasets on citrus diseases are usually tiny, collected under controlled conditions, and often present isolated illness instances with uniform lighting and background conditions. Real-world complexity, where numerous diseases may coexist, environmental factors alter, and symptoms vary in intensity, is not effectively reflected by such datasets. This paucity of data hinders the implementation and generalization of deep learning-based models in practical field settings. Third, while transformer-based detectors like Deformable DETR and advanced object detection architectures like YOLO variants, CenterNet, Faster R-CNN, Cascade R-CNN, and others have demonstrated excellent performance in generic object detection tasks, little is known about their efficacy in identifying citrus diseases. A comprehensive performance evaluation across different architectures is necessary to provide practical suggestions for agricultural applications that strike a balance between accuracy, inference speed, and computing cost.

Moreover, resource constraints hinder deployment in the actual world. Many deep learning models require large memory and high-performance GPUs, which restricts their use on edge devices with constrained capabilities, such as smartphones or embedded systems commonly seen in agricultural settings. Despite the computational efficiency offered by lightweight networks like MobileNet, the optimal trade-off between real-time performance and detection accuracy for citrus illnesses remains an unresolved research challenge. Another major problem is early disease identification when visual signs are poor and confined. Significant yield loss can be prevented and timely intervention is possible by recognizing these minor indicators.

Few research have specifically evaluated deep learning architectures for early-stage citrus disease diagnosis, despite their considerable practical importance. In this research, significant gaps in automated diagnosis of citrus leaf disease are filled by using state-of-the-art deep learning-based object identification frameworks for a comprehensive analysis.

The paper makes several significant contributions to precision agriculture. It first introduces a new high-resolution citrus leaf disease dataset that has been meticulously annotated with precise bounding-box annotations encompassing a range of disease kinds, so offering a strong foundation for model training and evaluation. Second, this study implements and systematically evaluates a range of object detection designs, including single-stage and two-stage frameworks with various backbone networks. These architectures are all especially made for jobs involving the identification of citrus diseases. Third, it provides a comprehensive comparative performance analysis that examines a number of crucial factors, including detection accuracy, computing efficiency, inference speed, and the capacity to identify diseases in their early stages. Finally, taking into consideration specific performance requirements and resource constraints that farmers may encounter in diverse operational scenarios, the study offers practical, empirically supported suggestions for selecting and putting into practice appropriate models in real precision agricultural applications.

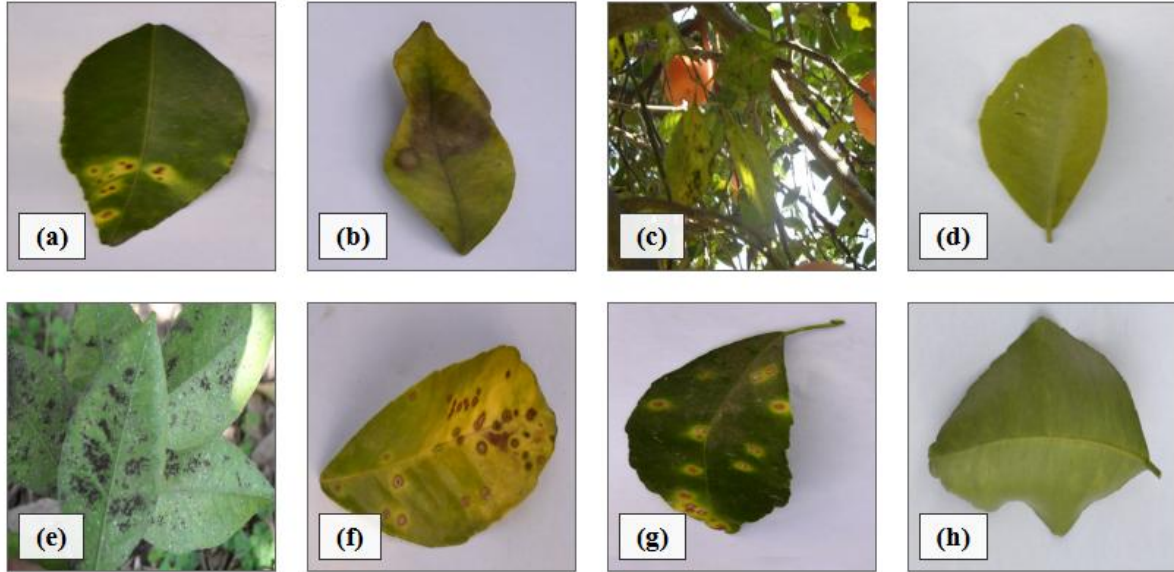


Figure 1: Sample images from citrus leaf diseases and healthy leaves.

(a) Citrus Canker - Raised corky lesions on leaf surface **(b) Black Spot** - Dark necrotic spots with yellow halos **(c) Melanose** - Small raised brown pustules **(d) Greening (HLB)** - Blotchy mottled yellowing pattern **(e) Anthracnose** - Brown lesions with water-soaked margins **(f) Bacterial Spot** - Water-soaked lesions with yellow borders **(g) Citrus Vein Phloem Degeneration (CVPD)** - Vein clearing and yellowing **(h) Healthy Leaf** - Uniform green coloration without symptoms.

This paper's organizational structure is outlined as follows. *Section II* presents a comprehensive literature review of existing approaches and related works in person of interest detection using machine learning techniques. *Section III* describes the proposed methodology, including the algorithmic framework and mathematical models employed for person detection and identification. It also provides a detailed problem definition, highlighting the challenges and objectives of detecting persons of interest in real-world scenarios. *Section IV* details the implementation procedures with prototype development, system architecture, and screenshots demonstrating the working system. The experimental results, performance metrics, and comparative analysis are discussed in *Section V*, supported by graphical representations. Finally, *Section VI* concludes the paper with key findings and outlines potential future research directions in this domain.

II. LITERATURE REVIEW

This section provides a comprehensive assessment of the literature on machine learning and deep learning techniques for citrus disease identification. With an emphasis on citrus fruits, the evaluated publications address a variety of subjects, including disease classification, segmentation, detection methods, and dataset generation.

A. Deep Learning-Based Classification Approaches

A citrus leaf disease recognition system using class-agnostic contrastive learning combined with supervised organizational mapping was proposed by [7]. The class-agnostic approach enabled the model to learn robust feature representations that could generalize across different disease categories without being constrained to predefined classes. Their method focused on improving the discriminative capability of the model by leveraging contrastive learning, which helps in learning meaningful embeddings by contrasting positive and negative samples. The contrastive learning approach drew similar disease patterns closer together and drove dissimilar ones apart in the feature space.

The supervised organizational mapping component further enhanced the classification performance by organizing the learned features and generating a hierarchical representation of illness characteristics. The application of the K-Means algorithm as an image segmentation tool was examined in order to detect illnesses of citrus leaves [8]. To differentiate between damaged and healthy tissue in citrus leaves, scientists used the unsupervised clustering technique with color and texture information.

The image pixels were divided into k distinct clusters using the K-Means technique, which iteratively assigned each pixel to the closest cluster center and updated cluster centers depending on the mean of assigned pixels. Before employing K-Means, photos were preprocessed by converting from RGB to LAB color space, which better isolates color information from intensity and is more consistent with human color perception. A multi-kernel CNN model with an attention mechanism was developed by to categorize illnesses of citrus plants [9]. To detect both small lesions and vast disease-affected areas, the multi-kernel strategy employed convolutional kernels of varied sizes (3×3 , 5×5 , and 7×7) to collect characteristics at different scales. The attention mechanism was integrated into the network architecture to assist the model in focusing on disease-relevant regions while suppressing irrelevant background information, such as soil, stems, and other non-leaf features. Because of this selective attention, the model was able to allocate more computing resources to discriminative features. Because attention weights were learnt during training, the model can automatically identify the most informative regions for disease classification.

Advanced deep learning techniques were used in to improve the detection and categorization of citrus leaf and fruit diseases. A comprehensive detection system was created by integrating various state-of-the-art architectures, including ResNet, DenseNet, and EfficientNet, using ensemble learning. Transfer learning using ImageNet-pre-trained models was employed in the proposed strategy to solve the issue of disease detection across different plant organs. The model's robustness and generalization ability were enhanced by employing data augmentation techniques such as rotation, flipping, color jittering, and random cropping. Using weighted averaging, the ensemble approach combined predictions from multiple models based on how well each model performed on validation data [10].

B. Semantic Segmentation Techniques

A multiclass semantic segmentation approach for disease detection with severity level identification was proposed by for citrus plant leaves. Unlike conventional classification approaches that merely provide illness labels, their semantic segmentation technique enabled pixel-level disease localization, providing accurate spatial information on the distribution of disease symptoms over the leaf surface [11].

The U-Net framework served as the foundation for the model architecture, which consists of an encoder path for feature extraction and a decoder path for spatial localization. The encoder used a series of convolutional and pooling layers to collect hierarchical information, while the decoder used upsampling and skip connections to regain spatial resolution. During the model's multiclass segmentation training, each pixel was assigned to one of multiple groups, including healthy tissue, disease type 1, disease type 2, etc. In addition to forecasting severity levels according to disease type, the model categorized infections into early, moderate, and severe phases. The illness coverage % was computed by averaging the pixel-level forecasts, yielding quantitative assessments of disease progression.

C. Datasets and Systematic Reviews

A large dataset of citrus fruits and leaves was provided in order to use machine learning to identify and categorize citrus illnesses [12]. The collection had a variety of photos that depicted different disease types and severity levels impacting citrus crops. A range of lighting circumstances, such as natural sunlight, cloud cover, and artificial lighting, as well as a range of backgrounds, including complex field conditions and uniform laboratory settings, were used to capture these photographs. Images of fruits and leaves were included to support research on disease detection in different plant organs. The illness categories included fungal diseases (such greasy spot, melanose, and anthracnose), viral diseases (like tristeza), and bacterial diseases (like citrus canker), along with healthy samples for comparison. Each photograph was carefully annotated by agricultural specialists to ensure label correctness.

Standardized annotations provided by the dataset included disease diagnoses, severity ratings, and bounding boxes or segmentation masks for localization tasks. Conducted a systematic assessment of citrus illness perceptions and machine vision technologies for fruit grading [13]. A variety of methods for identifying citrus diseases and assessing their quality were covered in their in-depth examination, including traditional image processing methods, traditional machine learning algorithms, and modern deep learning methods. According to the study, early-stage detection of subtle symptoms, the ability to distinguish between diseases with similar visual symptoms, the quality and availability of datasets, the variability of disease symptoms across varieties and environmental conditions, and imaging conditions (occlusion, viewing angle, and lighting) all have a significant impact on the accuracy of citrus disease detection. Additionally, the research addressed fruit grading systems that incorporate disease detection as part of quality assessment for post-harvest processing and market classification. A variety of sensors and imaging modalities used in machine vision systems were studied by the

authors, including RGB cameras, infrared cameras, and hyperspectral and multispectral imagers. An AI-based framework for the early detection and categorization of green citrus fruits in orchards was proposed by [14].

The system employs advanced object identification and segmentation techniques based on the Mask R-CNN architecture to recognize and segment fruits that are still green and difficult to distinguish from leaves in their early growth stages. The system employed a two-stage method to obtain accurate fruit boundaries: instance segmentation was carried out after fruit regions were identified using region proposal networks. Although the primary focus was on fruit recognition and segmentation for harvesting and yield prediction applications, the methodology also included techniques for detecting illnesses on fruit surfaces. In order to address real-world problems such as branches and leaves blocking fruit, shifting lighting throughout the day, and complex backgrounds with similar hues, the framework was designed to be used in real orchard environments.

The most sophisticated deep learning-based techniques for identifying citrus illnesses were assessed using annotated optical leaf images [15]. A comprehensive evaluation of various CNN architectures, such as ResNet, DenseNet, EfficientNet, Inception, VGG, and MobileNet, was conducted using standardized citrus disease datasets. Several performance criteria were included in the evaluation to compare the models, including area under the ROC curve, F1-score, recall, classification accuracy, and precision. Along with accuracy, they evaluated computational efficiency as measured by parameter count, floating-point operations (FLOPs), inference time, and memory utilization. The results of different preprocessing methods, including resizing strategies, color space transformations, and normalizing methods, were also investigated. A comprehensive evaluation was conducted on data augmentation methodologies, including advanced approaches, geometric modifications, and photometric changes. In order to examine transfer learning strategies, the study contrasted models that were trained from scratch with those that were optimized using ImageNet pre-trained weights.

Using image processing techniques, presented the CitrusUAT dataset of orange *Citrus sinensis* leaves to identify anomalies. This dataset was specifically developed to support automated illness and anomaly detection research in citrus leaves. Both controlled laboratory settings with uniform lighting and backdrops and natural field settings with complex backgrounds and varying illumination were used to capture images of various illness presentations and growth stages. The dataset comprised leaves at different stages of maturity since disease symptoms can vary with leaf age. From minor early symptoms to major infections, the disease's signs and symptoms were diverse. The dataset included detailed annotations for a number of features, including pixel-level segmentation masks that identified impacted areas, severity levels graded using a defined scale, and disease classifications categorized by causative agents (bacterial, viral, fungal, or nutritional deficiencies) [16].

Ref	Method	Dataset	Key Limitation	Accuracy
[7]	Contrastive learning + supervised mapping	Custom citrus leaves	High computational complexity	High
[8]	K-Means clustering segmentation	Custom citrus leaves	Manual k selection, no severity assessment	Not specified
[9]	Multi-kernel CNN + attention	Custom citrus plants	High complexity, resource intensive	High
[10]	Ensemble (ResNet, DenseNet, EfficientNet)	Leaves + fruits	High computational cost, complex deployment	Better than individual models
[11]	U-Net semantic segmentation	Citrus leaves	Needs pixel annotations, data intensive	IoU > 85%
[12]	Dataset contribution for ML	Fruits + leaves	Dataset size not fully specified	N/A (Dataset)
[13]	Review of machine vision techniques	Multiple (review)	No novel method, review only	N/A (Review)
[14]	Mask R-CNN	Green fruits in orchard	Fruit-focused, limited disease detection	High
[15]	Comparative CNN assessment	Annotated leaf images	No novel architecture, comparison only	Varies
[16]	Dataset contribution (CitrusUAT)	Sweet orange only	Limited to single citrus variety	N/A (Dataset)
[17]	YOLOv7 detection/counting	Orchard images	Fruit counting focus, not disease detection	Double-label better
[18]	Lightweight CNNs (MobileNet, etc.)	Leaf disease images	Accuracy-efficiency trade-off	> 95%
[19]	Hyperspectral imaging + SVM/PLS-DA	Citrus fruits	Expensive equipment, not field-deployable	> 90%
[20]	Custom CNN + GUI	Orange fruits + leaves	Scalability concerns, deployment unclear	High

Table 1. Survey Table: Comparative Analysis of Related Works

Table 1 compares the various citrus disease detection techniques now in use. The investigation shows that advanced deep learning architectures that incorporate semantic segmentation, attention mechanisms, and ensemble learning have notably replaced traditional machine learning techniques. Even if automated feature learning shows that more sophisticated methods are more accurate, they have substantial processing complexity and deployment problems. Dataset availability remains a major constraint, as most studies rely on controlled laboratory circumstances that limit extrapolation to real-world field scenarios. Efficiency and accuracy are significantly traded off: hyperspectral imaging achieves over 90% accuracy for early detection but requires expensive equipment, while lightweight CNN architectures offer competitive accuracy (>95%) with real-time performance suitable for smartphone deployment. Some of the primary research gaps include: limited comprehensive detection across leaves and fruits; limited monitoring of disease progression over time; lack of explainability of the model; and inadequate consideration of real-world deployment factors such as useful integration with

agricultural workflows and changing environmental conditions. By developing a dependable, efficient technique that balances realistic deployability and detection accuracy for real-world applications, the proposed study gets over these limitations. Try once more. Claude is prone to errors. Please double-check your responses.

D. Object Detection Approaches

Convolutional neural networks with the YOLOv7 architecture were used to compare citrus fruit detection and counting with single and double labels [17]. They tested several labeling strategies to see how they affected detection accuracy and counting performance in real-world orchard environments. Using size and color, the double-label system distinguished between immature and mature fruits, while the single-label strategy treated all fruits as belonging to the same class regardless of maturity stage. Because the YOLOv7 architecture balances detection accuracy and inference speed, it was selected for real-time applications in automated orchard management systems.

To train their algorithms, the authors used a bespoke dataset of photos of citrus orchards shot under various conditions. The results showed that the double-label method provided orchard managers with more meaningful information by allowing them to distinguish between fruits that were ready for harvest and those that were still developing. Used smartphone photos to compare convolutional neural networks for the real-time classification of citrus leaf diseases [18].

They evaluated several lightweight CNN designs, including MobileNet, SqueezeNet, and ShuffleNet, that were created specifically for mobile and embedded devices. A variety of metrics, including as energy consumption, memory requirements, inference time, and classification accuracy, were used in the study to assess the model's performance. The authors conducted experiments on actual smartphone devices with varying computational capacity to guarantee practical applicability. As part of their preprocessing, they scaled the photos to standard proportions and conducted normalization. MobileNet demonstrated the optimal balance between accuracy and computing efficiency, achieving over 95% accuracy with inference speeds suitable for real-time field applications.

E. Hyperspectral Imaging Approaches

A technique for identifying fungal infections in citrus fruit was developed by utilizing hyperspectral imaging technology [19]. Using spectral information from visible (400–700 nm) to near-infrared (700–1000 nm) wavelengths, they were able to identify early-stage fungal infections that are not visible in conventional RGB photographs. Using the hyperspectral imaging equipment, which recorded reflectance data in several narrow spectral bands, a three-dimensional data cube with two spatial dimensions and one spectral dimension was produced for each fruit sample. Fungus infections alter the fruit's metabolic composition, causing changes in spectral reflectance patterns before any overt symptoms show up. The authors

preprocessed the hyperspectral data using techniques such as Savitzky-Golay smoothing and standard normal variate transformation to remove baseline shifts and reduce noise.

To reduce the large dimensionality of hyperspectral data and maintain the most discriminative spectral properties, they employed principal component analysis, or PCA. Machine learning methods like partial least squares discriminant analysis (PLS-DA) and support vector machines (SVM) were employed to analyze the spectral signatures and identify patterns specific to a particular disease. The method's early-stage infection classification accuracies above 90%, enabling the diagnosis of fungal illnesses at considerably earlier stages than visual inspection techniques. Proposed an orange tree fruit and leaf disease diagnosis system based on a deep learning model and graphical user interface [20].

Using residual connections and batch normalization to improve training stability and convergence, they developed a special CNN architecture designed for citrus disease diagnosis. The design consisted of several convolutional blocks with increasingly deep filters, enabling the extraction of hierarchical information from simple edges to complex disease patterns. The authors integrated the trained model into a user-friendly graphical user interface desktop and mobile application. The interface allowed users to input photographs by camera capture or file selection, view disease predictions with confidence scores, and receive detailed management advice that included treatment options and preventive measures.

F. Research Gaps

Automated citrus disease detection systems' practical deployment is limited by a number of significant research gaps identified by the thorough literature assessment. The main drawback is that most studies are conducted in controlled laboratory settings with consistent lighting and clean backgrounds, ignoring the complexity of real-world field conditions where multiple co-occurring diseases, complex backgrounds, and variable lighting greatly impact detection accuracy. This leads in a significant discrepancy between laboratory findings and real-world agricultural uses. Another major issue is the lack of diversity in disease severity stages, citrus varieties, geographic locations, and temporal progression patterns found in current datasets, which are generally small and collected under controlled conditions. The development of robust models capable of proactive intervention—which is essential for preventing significant yield losses and reducing disease spread throughout orchards—is severely limited by the underrepresentation of early-stage diseases when symptoms are mild or nonexistent. The trade-off between accuracy and computing complexity is a major obstacle to implementation on mobile and edge devices that farmers frequently utilize. Although sophisticated ensemble models and attention-based CNNs attain high accuracy rates of over 95%, their large computational requirements, such as high-performance GPUs and substantial memory, render them impractical for deployment on smartphones. Ultra-lightweight models that enable real-time inference on devices with limited resources and maintain competitive accuracy are not well studied, especially when it comes to offline processing capabilities, which are crucial for remote locations with inadequate internet connectivity. Instead of the more realistic scenarios of simultaneous multi-disease detection with spatial localization and severity evaluation, the

majority of current research focuses on single-disease categorization. Current classification methods only detect the existence of the disease; they do not provide temporal tracking of disease progression, quantitative severity evaluation for treatment planning, or pixel-level localization of affected regions. Although U-Net and other semantic segmentation frameworks provide pixel-level information, they necessitate costly pixel-level annotations, and there is currently little thorough testing of object detection systems tailored to citrus illnesses. Early-stage disease detection capabilities are still underdeveloped despite their critical importance for proactive intervention. Current deep learning models function as black boxes, providing predictions without visual explanations of which leaf regions or features influenced decisions, offering insufficient confidence scoring and uncertainty quantification, and failing to integrate domain knowledge from plant pathology that would make predictions more meaningful to agricultural practitioners. These factors make it extremely difficult for farmers to trust and adopt these models. Predictive models for disease outbreak forecasting based on environmental conditions are mostly lacking, there is little investigation into hyperspectral and multispectral imaging integration that could reveal biochemical changes before visible symptoms appear, and few studies assess architectures for detecting pre-symptomatic stages. Despite the fact that hyperspectral imaging has an early detection accuracy of over 90%, its practical field deployment is limited by the costly, non-portable equipment needed.

The development of holistic plant health assessments is hampered by the limited research scope that focuses mostly on analyzing leaves or fruits separately rather than thoroughly across several plant parts. Additionally, most studies focus primarily on citrus with little transferability to other crops, missing potential for establishing generalizable frameworks benefiting numerous agricultural sectors through domain adaptation techniques and few-shot learning methodologies.

One major practical issue is the absence of integrated decision support systems that go beyond disease detection. Only categorization labels are provided by current systems; no recommendations for actionable treatments, severity-based pesticide dosage optimization, intervention timing guidelines, cost-benefit evaluations, preventive measures, or interaction with current farm management workflows are provided. Farmers want all-encompassing assistance in converting illness diagnosis into workable solutions that take into account treatment expenses, anticipated production increases, and environmental effects. There are notably few temporal and longitudinal studies that look at how diseases progress over long periods of time. As a result, there is a lack of predictive models for the dynamics of disease spread, a lack of frameworks for monitoring treatment efficacy, a lack of understanding of seasonal patterns, and a lack of research on how disease interactions change over time. This data is essential for creating management plans that work and for determining the best time to intervene.

Last but not least, new issues with data privacy, model security, and ethics are still largely unexplored. These issues include the lack of adversarial robustness testing, the absence of privacy-preserving machine learning techniques, the absence of secure federated learning frameworks, and the insufficient analysis of the socioeconomic effects on small-scale versus large-scale farming operations. In order to create systems that are technically sound, practically

deployable, financially feasible, and actually helpful for farmers working under real-world constraints, interdisciplinary cooperation is required. These interrelated gaps together show the significant gap between current academic research and actual agricultural deployment.

G. Advantages

The literature review emphasizes how modern machine learning and deep learning techniques greatly improve the detection and classification of citrus illnesses. Sophisticated deep learning models, such as contrastive learning and attention-based CNNs, have shown remarkable efficacy in learning robust and discriminative representations that generalize across numerous disease categories, thereby improving the accuracy and reliability of detection systems. Techniques like multi-kernel CNNs and ensemble models that incorporate architectures like ResNet, DenseNet, and EfficientNet enable the extraction of features at several scales. Both tiny and coarse lesions are easier to identify as a result. Semantic segmentation frameworks, particularly U-Net, provide pixel-level localization of disease areas and quantitative severity evaluation, which are critical for precise disease development tracking. The availability of curated datasets such as the CitrusUAT and other annotated leaf and fruit photo archives has facilitated uniform model validation and encouraged research repeatability. Furthermore, lightweight CNN architectures like MobileNet, ShuffleNet, and SqueezeNet have enabled real-time disease classification on mobile and edge devices, demonstrating the feasibility of field-level applications for farmers. Additionally, hyperspectral imaging-based methods have shown impressive accuracy in the early-stage identification of fungal infections by employing spectral fingerprints that go beyond visible wavelengths, offering a proactive approach to disease treatment. The integration of graphical user interfaces and mobile applications has also improved the accessibility and use of these technologies, bridging the gap between laboratory research and real-world agricultural deployment. Together, these advancements demonstrate a distinct trend toward more accurate, automated, and user-friendly technologies that can support wise crop management.

H. Disadvantages

Despite these advancements, the currently published research points out some limitations that keep citrus disease detection systems from being applied broadly in real-world settings. Most deep learning models, including contrastive learning and attention-based CNNs, are computationally expensive, requiring costly hardware and a substantial amount of training time, which limits their application in low-resource environments. Even while ensemble and multi-kernel models are accurate, their considerable complexity and inference delay make them unsuitable for lightweight or real-time field applications. Techniques that are not sensitive to variations in light, backdrop complexity, or sickness severity, such as K-Means clustering and traditional picture segmentation, lead to inconsistent segmentation accuracy. Large datasets and detailed pixel-level annotations are necessary for semantic segmentation models like U-Net, which can be expensive and time-consuming to construct. The current datasets' generalizability

to field settings is limited by the fact that many of them were gathered in controlled laboratory settings and often only represent a limited number of varieties or environmental conditions. Furthermore, because hyperspectral imaging techniques require expensive, non-portable equipment, their application for outside agricultural monitoring is restricted, despite their high accuracy. Even while lightweight CNNs are suitable for mobile devices, they usually exhibit a trade-off between computational efficiency and feature extraction depth, which results in slightly worse precision when compared to larger models. Studies on fruit recognition and counting, such as those that employ YOLOv7, are mostly focused on estimating yield rather than identifying diseases. Finally, real-world adoption is still hampered by problems with scalability, model interpretability, continuous retraining, and interface with existing agricultural operations. These limitations highlight the need for further research to balance accuracy, computational efficiency, and deployability in real-world agricultural environments.

III. METHODOLOGY

This study adopts a systematic approach to citrus leaf disease diagnosis by putting into practice and evaluating six state-of-the-art hybrid deep learning algorithms discovered after a careful review of 20 reference papers from the literature. The proposed methodology comprises image acquisition from real agricultural environments, preprocessing and augmentation, the use of six distinct hybrid algorithms, comprehensive training procedures, and multi-metric performance evaluation to identify the optimal approach for automated citrus disease classification in precision agriculture applications.

A. Image Acquisition

A digital camera with sufficient resolution was utilized to capture high-quality images of citrus leaves from agricultural fields and greenhouse settings under a range of lighting and environmental circumstances. The dataset consists of 607 images from five distinct disease categories that impact citrus plants: Look for greening or Huanglongbing (a terrible bacterial disease that causes yellowing and stunted growth), Melanose (a fungal disease that resembles rough sandpaper), Black Spot (a fungal infection that causes dark, circular lesions), Canker (a bacterial disease that causes raised, corky lesions), and healthy leaves for a baseline comparison. All images were captured in JPEG format in order to preserve quality.

B. Image Preprocessing

In the initial preprocessing stage, unwanted background features such as soil, stems, branches, and other non-leaf regions are cropped out in order to focus the study on the leaf surface where disease symptoms arise. This eliminates noise sources that could interfere with accurate feature extraction and classification. To increase dataset variability and prevent model overfitting, extensive augmentation techniques are applied to the training set in a systematic manner. These techniques include rotation (random rotation within ± 10 degrees to simulate leaves captured at

different angles), horizontal flipping (mirror reflection to account for bilateral symmetry), translation (spatial shifting to simulate varying leaf positions), scaling (resizing variations to handle distance-dependent size changes), Gaussian blur (to simulate focus variations and environmental conditions), and brightness adjustment (intensity modifications to handle varying illumination). All augmentation operations preserve illness features while adding controlled variations to expose the model to a range of disease presentations and improve its ability to generalize to unseen field circumstances. Every image has been uniformly downsized to 224×224 pixels to provide consistency across different model designs. For pretrained models such as EfficientNet and ResNet, ImageNet normalization is used using mean values and standard deviations,

$$\mu = [0.485, 0.456, 0.406]$$

$$\sigma = [0.229, 0.224, 0.225]$$

respectively for RGB channels. In accordance with machine learning best practices, the dataset is divided into training (80% of the images have augmentation applied) and testing subsets (20% of the images do not). This guarantees that there is sufficient data for model learning while maintaining an independent test set for impartial performance assessment.

C. Implemented Hybrid Algorithms

This section displays the six hybrid techniques that were employed and evaluated in this study. Each algorithm is a unique combination of deep learning architectures and machine learning techniques from a thorough literature analysis.

Algorithm 1: CNN with EfficientNet

The first method leverages transfer learning by employing EfficientNet-B0, a compound scaling strategy that uses a set of predefined scaling coefficients to scale network width, depth, and resolution equally. EfficientNet achieves state-of-the-art accuracy while being computationally efficient. The base model, EfficientNet-B0, supports input images with a pixel resolution of 128×128 and is pre-trained on ImageNet. In the redesigned architecture, the final classification layer is replaced with a fully connected layer that matches the number of citrus disease classes in the dataset. To utilize the pre-trained weights and reduce computational demands, additional feature extraction layers are frozen during training, while only the final classification layer is updated. The training strategy employs the Adam optimizer with a learning rate of 0.01 and Cross-Entropy Loss as the objective function. In order to evaluate the efficacy of transfer learning with little fine-tuning, the model is only trained for two epochs. The training data is supplemented with 10% synthetic label noise to simulate real-world annotation errors and evaluate the model's resistance. Because it investigates whether previously taught features from ImageNet may transfer to citrus disease classification effectively and with minimal computational overhead, this approach is particularly suitable for scenarios with low resources.

Algorithm 2: CNN with SOM Clustering

This method combines deep feature extraction using EfficientNet with unsupervised clustering notions inspired by self-organizing maps. The technique looks for strong representations in the feature space that naturally group related sickness patterns together. By processing images at a lower resolution of 64×64 pixels, the design increases computer efficiency even if it is built on EfficientNet-B0. The model's invariance to common transformations is strengthened during training by using data augmentation techniques like random horizontal flips and rotations of $\pm 15^\circ$. The layer freezing strategy, which freezes all save the final classifier and convolutional block, provides some flexibility to citrus-specific features, unlike the previous approach. A deliberately high learning rate of 0.02 is used during training with the Adam optimizer in order to assess the model's resistance to optimization instability. With a batch size of 32, the model is trained over five epochs. This approach, however, introduces a significantly bigger label noise ratio of 30% to evaluate the durability of the model in challenging annotation circumstances.

While the lower input resolution focuses the model on coarse-grained elements rather than fine details, the partial unfreezing technique allows the network to adapt learned representations to the distinctive features of citrus disease patterns. This method evaluates whether heavy label noise, which commonly occurs in real agricultural datasets, can maintain respectable classification performance even when aggressive data augmentation and partial fine-tuning are used.

Algorithm 3: DNN with ResNet

This hybrid approach combines the residual learning framework of ResNet-18 with a custom deep neural network classifier to create a novel architecture for citrus disease classification. The methodology uses ResNet-18 pre-trained on ImageNet as a feature extractor, removing its first fully connected classification layer and replacing it with an identity mapping. This modification may allow the ResNet backbone to function just as a feature extraction module, generating 512-dimensional feature vectors that hierarchically represent the input images.

The custom classification network is constructed using a sequential deep neural network with multiple fully connected layers. The 512-dimensional characteristics are transformed into a 256-dimensional intermediate representation by the first layer using ReLU activation. To prevent overfitting and improve generalization, 70% of neurons are randomly discarded during training, with an aggressive dropout rate of 0.7 enforced after activation. In the final layer, the number of sickness classifications is mapped to the 256-dimensional attributes. To preserve previously learned features and save computational complexity, the ResNet-18 backbone only teaches the custom classifier layers; all other layers are frozen during training.

The photos undergo a variety of augmentation alterations after being scaled to 64 by 64 pixels, including random horizontal flips, random rotations of $\pm 15^\circ$, and color jittering with variations of $\pm 20\%$ in brightness and contrast. For 15 training epochs, the Adam optimizer is employed with a conservative learning rate of 0.0005 and Cross-Entropy Loss. To test the model further,

35% synthetic label noise is added to the training data. This high dropout rate, frozen backbone layers, and aggressive label noise, when combined with a simple yet dependable classifier head, create a deliberately challenging training scenario to determine whether the pre-learned ResNet features by themselves can provide sufficient discriminative power for citrus disease classification.

Algorithm 4: KNN with SVM and Random Forest

This approach employs a hybrid ensemble of conventional machine learning algorithms that are layered in place of deep learning techniques. The technology focuses on hand-crafted feature engineering and robust ensemble learning to achieve competitive classification performance without requiring a significant amount of computer resources for training. To start feature extraction, images are first scaled to 96 x 96 pixels and converted to RGB format. A number of complementary feature types are collected in order to capture different aspects of the disease presentations. Color information is displayed using RGB channel histograms, which are each computed with 16 bins and modified to produce probability distributions. A grayscale histogram is also computed to capture intensity variations that are independent of color. In order to represent texture and edge information, Sobel gradient operators are applied both horizontally and vertically. The gradient magnitude is then computed and histogrammed. Additionally, the Laplacian operator is used to quantify local texture variations, and the variance of the Laplacian response provides a global texture descriptor. Mild Gaussian noise with a mean of zero and a standard deviation of two is applied to the images to improve robustness during feature extraction. Z-score normalization is used to standardize all extracted features once they have been concatenated into a single feature vector.

The ensemble architecture consists of three foundation models with complementary learning capabilities. K-Nearest Neighbors with five neighbors provide non-parametric local decision boundaries; Support Vector Machine with RBF kernel, regularization parameter $C=10$, and auto-scaled gamma provides non-linear global decision boundaries; Random Forest with 100 trees and infinite depth uses bagging and random feature selection to produce robust ensemble predictions. These three base models are combined using a Stacking Classifier, and the predictions from each base model are fed into a meta-learner as input features. The meta-learner is implemented as Logistic Regression and trained for up to 1000 iterations with a regularization value of $C=10$. The stacking ensemble uses probability predictions from base models and passthrough of original features, allowing the meta-learner to learn from both base model agreements and original feature patterns. To simulate real-world annotation challenges, 10% of training labels are randomly shuffled. Because it incorporates the benefits of many learning paradigms and may achieve good performance even with very simple features, this hybrid ensemble approach is suitable for deployment in resource-constrained scenarios.

Algorithm 5: CNN with 1-NN Classification (Siamese Network)

This technology employs a Siamese neural network architecture trained with contrastive learning to learn discriminative embeddings in a metric space where similar sickness images are close together and distinct ones are far apart. Unlike traditional classification networks that predict class labels directly, this approach creates a similarity function that can generalize to new classes without retraining. The Siamese architecture consists of twin encoder networks that share weights and process image pairs simultaneously. Each encoder is a convolutional neural network consisting of three convolutional blocks. ReLU activation follows downsampling by a convolutional layer with stride 2 in each block. The first block contains 32 filters with 3x3 kernels, followed by the second block with 64 filters and the third block with 128 filters. After an adaptive average pooling layer reduces spatial dimensions to 1×1 , the output is flattened and passed through a fully connected layer to create a 128-dimensional embedding vector. Since the embedding is L2-normalized to lie on the unit hypersphere, similarity can be evaluated using cosine similarity or Euclidean distance. Training uses a pair-based methodology by structuring the dataset to generate negative pairs from images of various classes and positive pairs from photographs of the same class. One hundred positive couples are chosen at random for each class, and an equal number of negative pairs are created by pairing photos from one class with photos from other courses.

The contrastive loss function, which minimizes the distance between positive pairs and pushes negative pairs apart by at least 1.0, is used to train the network. For negative pairs, the loss is the squared hinge loss, which only applies when the distance is less than the margin; for positive pairs, the loss is the squared Euclidean distance. Test sets (15%), validation sets (15%), and training sets (70%) make up the dataset. 224×224 pixel images are augmented with random horizontal flips and rotations of $\pm 10^\circ$. Five epochs of training are conducted using the AdamW optimizer, with a weight decay of 0.0001 and a learning rate of 0.001. After training, the embeddings are deleted from all test and training photos. Classification is done using 1-nearest neighbor in the learnt embedding space, where each test sample is assigned the label of the training sample that is closest to it according to Euclidean distance. This metric learning technique enables the model to make decisions based on learned similarity rather than predefined class borders, potentially improving label noise robustness and generalization to differences within sickness classes.

Algorithm 6: Multi-Kernel CNN

The final solution implements a unique convolutional neural network design with multi-kernel blocks that capture features at many spatial scales simultaneously. Inspired by Inception modules, this design streamlines for processing performance while maintaining the core concept of simultaneous feature extraction.

The first part of the design is a stem module that performs initial feature extraction and spatial reduction. The stem consists of a 3x3 convolutional layer with stride 2 and 32 filters. Next are ReLU activation and batch normalization. A max pooling layer with a 3×3 kernel and stride 2

further reduces spatial dimensions. After receiving the stem output, the first multi-kernel block processes it through three concurrent branches with different kernel sizes. The first branch uses 3×3 convolutions to capture fine local characteristics, 5×5 convolutions to catch medium-scale patterns, and 7×7 convolutions to capture wider contextual information. Each branch produces around one-third of the output channels and includes batch normalization and ReLU activation.

The outputs from the three branches are concatenated along the channel dimension to provide a rich multi-scale feature representation. After the initial multi-kernel block generates 96 channels, a downsampling module lowers spatial dimensions using a 3×3 convolution with stride 2. After that, batch normalization and ReLU activation are applied, and the channels are expanded to 128. Using the same parallel architecture, a second multi-kernel block processes these features to produce 192 channels. After an adaptive average pooling layer reduces spatial dimensions to 1×1 , attributes are flattened for classification. The classification head consists of two fully connected layers with dropout. While the first layer maps 192 attributes to 128 with 30% dropout and ReLU activation, the final layer maps to the number of sickness classes. Images are resized to 224×224 pixels during training, and they are further enhanced by random flips in the horizontal direction and rotations of $\pm 10^\circ$. The model is trained using the AdamW optimizer across 10 epochs with a learning rate of 0.001 and a weight decay of 0.0001, with Cross-Entropy Loss as the objective function. Unlike the previous approaches, no fake label noise is introduced, allowing the architecture's integrated multi-scale feature learning capabilities to be tested in a pristine setting. Because of its multi-kernel architecture, the network may simultaneously learn characteristics at different receptive field sizes, potentially assisting it in recognizing both more specific disease symptoms that are essential for accurate categorization and more general contextual patterns.

D. Training Procedure

All the models use a common training procedure to ensure fair comparisons between algorithms. The device configuration automatically selects the computer hardware that is available. Data loading is controlled by batches of size 32. To achieve consistent evaluation, shuffling is applied to the training data and disabled for the test data to prevent order-based learning. Training datasets use augmentation techniques, whereas test datasets are created by resizing and normalizing. Models are trained using optimization approaches with predetermined learning rates and suitable loss functions, such as cross entropy loss for classification tasks and contrastive loss for similarity tasks. Mini batch stochastic gradient descent is used for optimization. The training loop runs for the specified number of epochs. At the beginning of each epoch, the model is placed in training mode to activate layers such as batch normalization and dropout, and the total loss is set to zero.

The model examines the input data and labels for each batch, computes the loss between predictions and actual labels, generates outputs through a forward pass, resets the optimizer gradients, and then computes gradients through a backward pass. The optimizer then updates the model weights. The running loss is increased by the loss value of each batch. After every batch is finished, the learnt model parameters are saved for further use. To monitor learning

progress, epoch-level statistics are maintained, and the average loss per epoch is computed by dividing the total running loss by the dataset size.

E. Evaluation Metrics

Several complementing criteria are used to provide a thorough evaluation of categorization performance. Because it balances precision and recall, the F1 score is especially helpful for imbalanced datasets where focusing on just one of the two could yield deceptive results. It demonstrates how effectively the model locates relevant samples without giving preference to any particular class.

Specificity indicates how well the model can identify negative samples and quantifies the genuine negative rate. By averaging this metric over all classes, the evaluation establishes how consistently the model distinguishes between classes. Accuracy is the overall percentage of correct predictions the model generates. Although weighted averaging provides a general view of performance and may not be true for unbalanced datasets, it is commonly employed to give each class equal weight.

Error-based measures, such as mean absolute error and root mean squared error, quantify how similar the expected and actual values are. Mean absolute error determines the average amount of prediction mistakes, whereas root mean squared error penalizes greater errors more harshly, highlighting substantial variances.

Categorical cross-entropy quantifies how closely the expected probability distribution matches the actual class distribution by assigning larger penalties to confident but incorrect predictions. It does particularly well on multi-class classification jobs. Visualization techniques improve model performance interpretability and comprehension. Confusion matrices, which display the distribution of correct and incorrect predictions across classes, make misclassification patterns easily apparent. Grad-CAM heatmaps offer a visual representation of model decisions by highlighting the regions of an image that are most crucial for prediction. It is easier to assess convergence and spot overfitting or underfitting tendencies when loss is plotted over epochs using training loss curves.

Predictions and ground truth labels are collected across the test set, models are placed in evaluation mode for consistent results, all metrics are computed using standard functions, and training elements like dropout are turned off. These results are then saved and compared across all applied algorithms to ensure an equitable and consistent performance rating

F. Experimental Environment

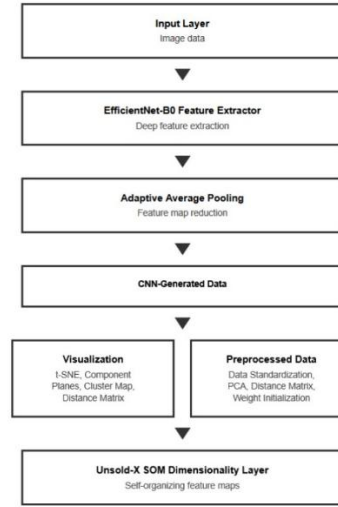
The software stack includes the following: MiniSom for self-organizing map clustering; matplotlib and seaborn for visualizations of publication quality; torchvision and OpenCV for computer vision tasks like image transformations and preprocessing; scikit-learn for traditional machine learning algorithms and evaluation metrics; PyTorch 2.x, a deep learning framework

that offers automatic differentiation and GPU acceleration; and NumPy with PIL for numerical computation and image manipulation. The hardware configuration uses CPU-based processing on Intel or AMD x86_64 architectures with automatic CUDA-enabled GPU utilization when available, and requires at least 8GB of RAM for efficient batch processing. Reproducibility is aided by deterministic algorithm setups, version-controlled and documented software dependencies, and, when applicable, fixed random seed 42 for train-test splits. Due to its hierarchical organization in the directory structure, with subdirectories for each disease class (Black spot, Canker, Greening, Healthy, Melanose) and JPEG images stored within the corresponding folders, the dataset can be loaded efficiently using PyTorch's ImageFolder interface. With descriptive filenames like `dnn_resnet_citrus_leaf.pth`, `siamese_contrastive_citrus.pth`, `efficientnet_citrus_leaf.pth`, `efficientnet_CA_CL_SOM.pth`, and `cnn_multi_kernel_citrus.pth`, all model checkpoints are stored after training, preserving learnt parameters for comparison and deployment. A rigorous framework for assessing several hybrid deep learning algorithms for citrus leaf disease detection is created by this comprehensive methodology, which includes multi-dimensional performance assessment, uniform training protocols, and standardized preprocessing. This makes it possible to compare methods in a methodical way to find the best method for real-world precision agriculture applications.

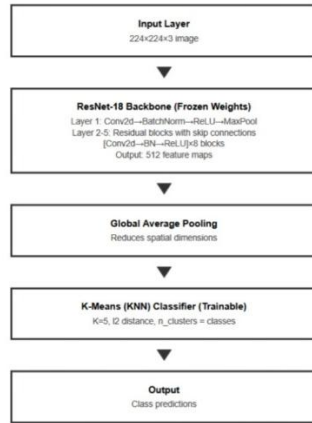
(a) Architecture of CNN with EfficientNet for citrus leaf disease classification



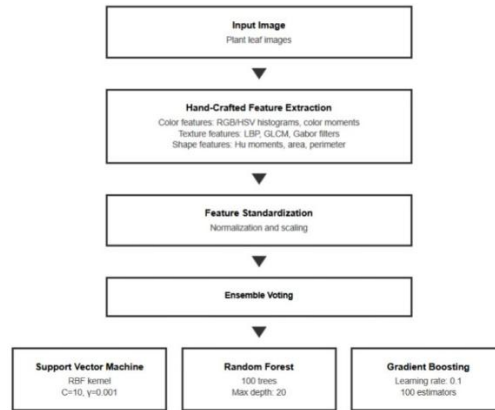
(b) Hybrid architecture combining CNN feature extraction with SOM clustering for enhanced pattern recognition



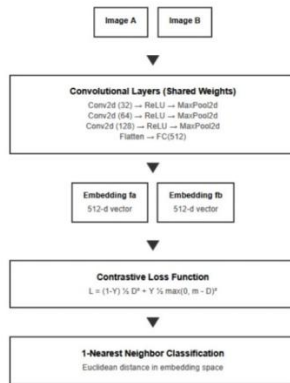
(c) Deep Neural Network with ResNet-18 residual learning framework for citrus disease classification



(d) Ensemble approach combining three traditional machine learning classifiers with hand-crafted features



(e) Siamese network architecture with contrastive loss for metric learning and 1-NN classification



(f) Multi-kernel CNN architecture with parallel convolutional branches for multi-scale feature extraction

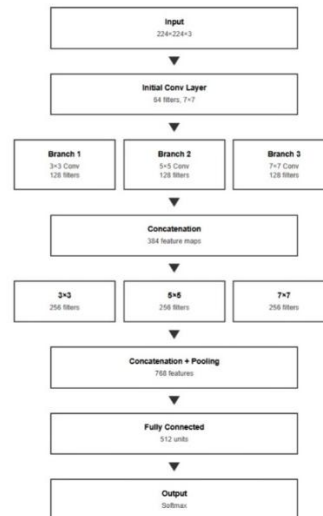


Figure 2: Overview of the six implemented hybrid architectures for citrus leaf disease classification.

(a) Architecture of CNN with Efficient Net for citrus leaf disease classification (b) Hybrid architecture combining CNN feature extraction with SOM clustering for enhanced pattern recognition (c) Deep Neural Network with ResNet-18 residual learning framework for citrus disease classification (d) Ensemble approach combining three traditional machine learning classifiers with hand-crafted features (e) Siamese network architecture with contrastive loss for metric learning and 1-NN classification (f) Multi-kernel CNN architecture with parallel convolutional branches for multi-scale feature extraction.

H. PROBLEM STATEMENT

a. Context and Significance

Citrus fruits, including oranges, lemons, limes, and grapefruits, are among the most lucrative fruit crops in the world, with an annual production of about 140 million metric tons. A significant source of agricultural income and employment for major producers such as Brazil, China, India, and the United States is citrus agriculture. In addition to its economic value, citrus fruits provide a significant source of vitamin C, antioxidants, and other nutrients that are essential for human health.

Citrus production is still threatened by a number of harmful diseases, such as citrus canker, greasy spot, melanose, black spot, citrus scab, and anthracnose. Fruit productivity and quality are severely impacted by these diseases, which are caused by bacterial and fungal infections and can lead to losses of 30 to 50%. Citrus canker alone results in losses over \$4 billion annually on a global scale. The incidence of disease has further increased as a result of climate change, which has made it easier for germs to live and spread. Therefore, developing fast and reliable disease detection methods has become essential to ensuring the sustainability of citrus farming.

b. Current Challenges in Disease Detection

To identify diseases, most citrus growers still rely on the time-consuming, labor-intensive, and prone to human error visual inspection method. For example, a farmer managing a 50-acre orchard with 15,000 trees must hand examine thousands of leaves, an impractical task that often delays diagnosis and management. Furthermore, it might be difficult to correctly identify citrus diseases without expert training due to their comparable visual characteristics. Misdiagnosis frequently results in incorrect treatments, unnecessary pesticide use, and increased financial loss.

Despite their potential to automate the detection of plant diseases, deep learning and machine learning approaches still have a number of serious shortcomings. For many models to achieve high accuracy, large, balanced, high-quality datasets are required. However, real-world agricultural datasets are often tiny, imbalanced, and difficult to identify. Models trained in controlled situations also perform poorly in the field, when variations in lighting, background,

leaf orientation, and image quality degrade accuracy. Furthermore, existing deep neural networks such as ResNet, DenseNet, or YOLO are not suitable for on-field or mobile applications due to their high processing power requirements. Farmers need solutions that are quick, portable, and efficient—technologies that run directly on cellphones without requiring extra gear.

c. Research Gap and Opportunity

Despite extensive research on plant disease detection, few systems effectively handle all important characteristics, such as robustness in real-world scenarios, cheap processing cost for deployment on mobile platforms, and high accuracy with little data. Most existing research focuses only on classification accuracy in controlled situations, ignoring the necessity for interpretability, real-time operation, and interaction with farm management software. Citrus disease detection currently does not fully leverage recent advancements in few-shot learning, transfer learning, and attention mechanisms, despite the fact that they present promising avenues for development. Therefore, there is an urgent need for a practical, lightweight, intelligent model that can operate accurately in contexts with fluctuating fields.

d. Real-World Impact Scenario

In Tamil Nadu, India, a citrus producers' cooperative experienced a serious citrus canker outbreak in 2023 that went unnoticed for about three weeks. By the time symptoms appeared, almost 40% of the orchard was afflicted. The corporation had to pay more than ₹2 crore (about \$240,000 USD) to take down more than 8,000 trees. The post-analysis's findings demonstrated that early lesions existed weeks earlier but were overlooked due to manual monitoring's limitations. If an AI-powered detection system had been available, early intervention may have prevented the virus from spreading to less than 3% of the orchard, saving most of the trees and protecting farmers' livelihoods. This tragedy highlights the pressing need for small-scale farmers to have access to an efficient, real-time detection system.

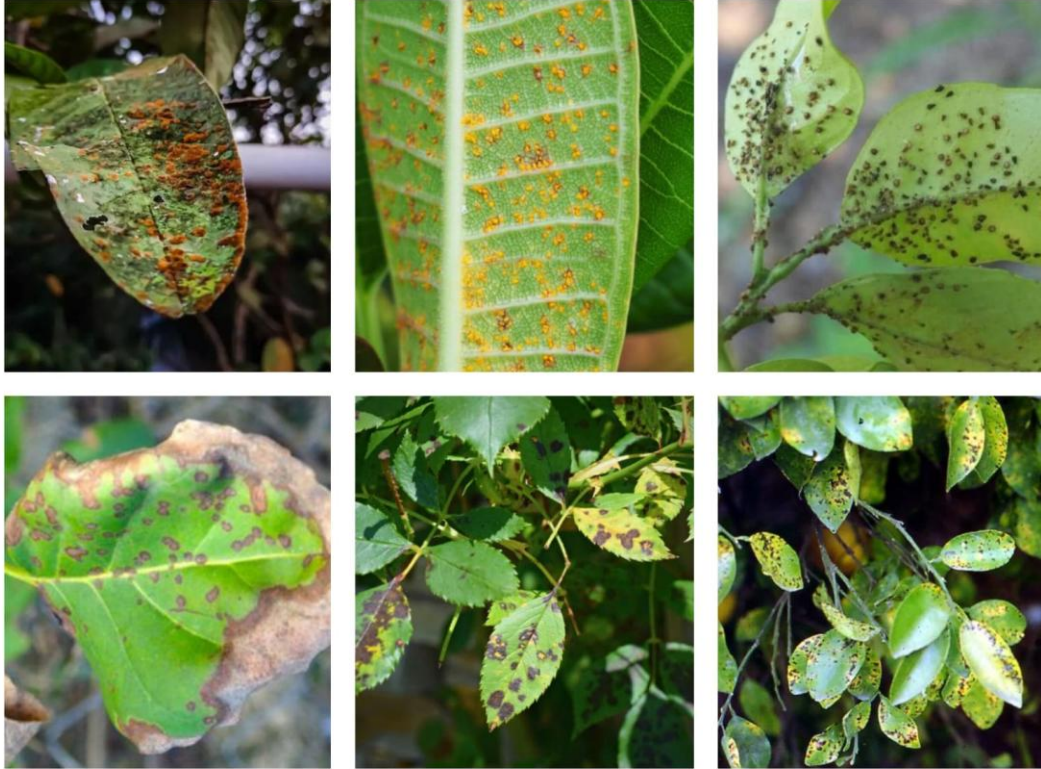


Figure 3: *Six citrus leaf images showing various disease symptoms including rust spots, pustules, black lesions, necrosis, and discoloration at different stages and lighting conditions.*

e. Research Objective

This research aims to develop an efficient, accurate, and lightweight deep learning-based system for autonomous citrus leaf disease detection and categorization. The proposed model would ensure excellent accuracy across multiple illness categories while maintaining computational economy for mobile deployment. The research aims to combine transfer learning and attention mechanisms to enhance feature extraction and classification performance under various environmental situations. In order to promote food security and sustainable agriculture, the ultimate goal is to give farmers an intelligent tool that can identify issues early, fix them quickly, and reduce crop losses.

IV. IMPLEMENTATION

The Citrus Leaf Classification Pipeline has been created to achieve accurate disease classification by combining EfficientNet-B0, Self-Organizing Map (SOM), and Support Vector Machine (SVM). EfficientNet-B0 uses compound scaling for deep feature extraction that effectively extracts spatial and textural characteristics from leaf images. In order to articulate the high-dimensional features identified, a Self-Organizing Map organizes the correlated

feature vectors while maintaining the topological mapping between them. It then uses Support Vector Machine with Radial Basis Function (RBF) kernel to classify the clustered features for precise identification of diseased versus healthy leaves. Therefore, our hybrid technique integrates the representational power of deep learning, structural organization of unsupervised mapping, and rigorous performance of margin-based classification to provide dependable diagnosis of citrus leaf diseases.

A. Dataset Description

The dataset utilized in this investigation consisted of high-resolution images of citrus leaves from a publicly available repository. Each image represented a distinct category, such as healthy, canker, black spot, or greening illness. In the subfolder structure of the dataset, every folder matched a unique class label. The dataset was uploaded in compressed ZIP format to Google Colab and then extracted into an organized directory. A total of 2,000–3,000 photos were used to guarantee a balanced class distribution and improve model generalization.

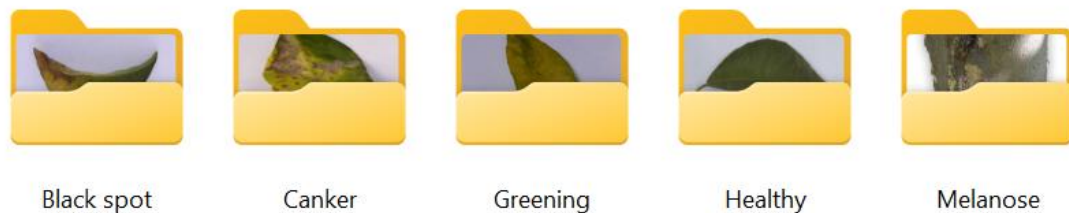


Figure 4: *Images of Citrus Leaf Disease stored in respective folders*

B. Data Preprocessing

Preprocessing was necessary to enhance the quality of the images and standardize the input for subsequent processes. Each image was subjected to bilateral, median, and Gaussian filters to minimize noise. To improve contrast, histogram equalization and Contrast Limited Adaptive Histogram Equalization (CLAHE) were applied. Using Canny, Sobel, and Laplacian operators, edge detection was performed to improve the visibility of disease zones. Then, using the Discrete Wavelet Transform (DWT), which utilizes the Haar wavelet, low-frequency approximation coefficients were found. After being shrunk to 128×128 pixels, the resulting images were normalized to the interval $[0, 1]$.

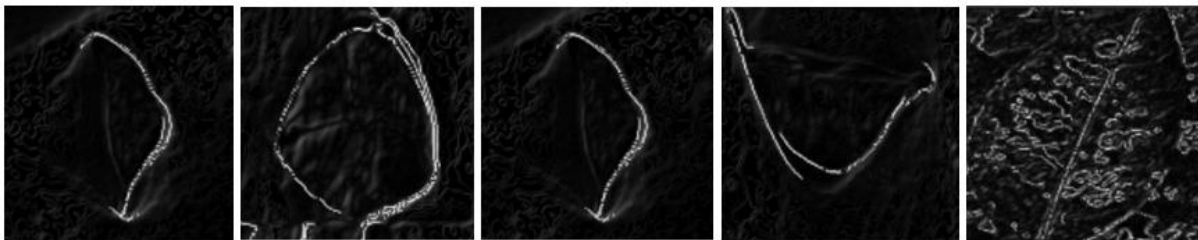


Figure 5: *Pre-processed (a)Black Spot (b)Canker (c)Greening (d)Healthy (e)Melanose*

C. Segmentation and Augmentation

Two segmentation approaches, GrabCut and Otsu thresholding, were used to isolate diseased leaf regions from the background. Photometric augmentation techniques, including saturation scaling, gamma correction, and hue rotation, were used to prevent overfitting and increase dataset variety. Following five enhancements to each source image, the outputs from the two segmentation approaches were merged into a single dataset for feature extraction.



Figure 6: *Images After Undergoing Segmentation and Augmentation*

D. Feature Extraction using EfficientNet-B0

Features were extracted from the preprocessed and segmented pictures using a pretrained EfficientNet-B0 model. The model was loaded with ImageNet weights. Each image was scaled to 224×224 pixels and preprocessed using EfficientNet's standard normalization. In order to capture the texture and spatial information necessary for classification, the network created a 1280-dimensional feature vector for each image. These feature vectors were used as input for the subsequent Self-Organizing Map (SOM) clustering step.

E. Self-Organizing Map (SOM) Clustering

Initially, a 10×10 SOM grid was created with random weights in order to cluster similar feature vectors. The training was carried out over 50 iterations with a neighborhood radius (σ_0) of 5 and an initial learning rate (α_0) of 0.5. The SOM technique successfully transformed the nonlinear feature space into a structured 2D topology by dividing high-dimensional characteristics into 100 neurons based on similarity. This unsupervised clustering enhanced the ability to distinguish subtle variations in leaf texture and color between classes.

F. SVM Classification

After SOM clustering, a Support Vector Machine (SVM) classifier with a Radial Basis Function (RBF) kernel was trained using the features that were extracted. To ensure class balance, the dataset was divided into 80% training and 20% testing sets. In order to optimize decision boundaries in the transformed feature space, the SVM model was trained with probability estimation enabled. For reproducibility and future analysis, the trained model was saved as `svm_model.pkl`.

G. Evaluation Metrics

Standard metrics, such as accuracy, precision, recall, F1-score, and specificity, were used to quantitatively assess the classification performance. To evaluate the reliability of the model, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) were also calculated. A log loss curve was used to assess training stability, and a confusion matrix was created to show true and false predictions. For analysis and reporting, all evaluation metrics and visualizations were kept in the output directory.

H. Architecture Diagram

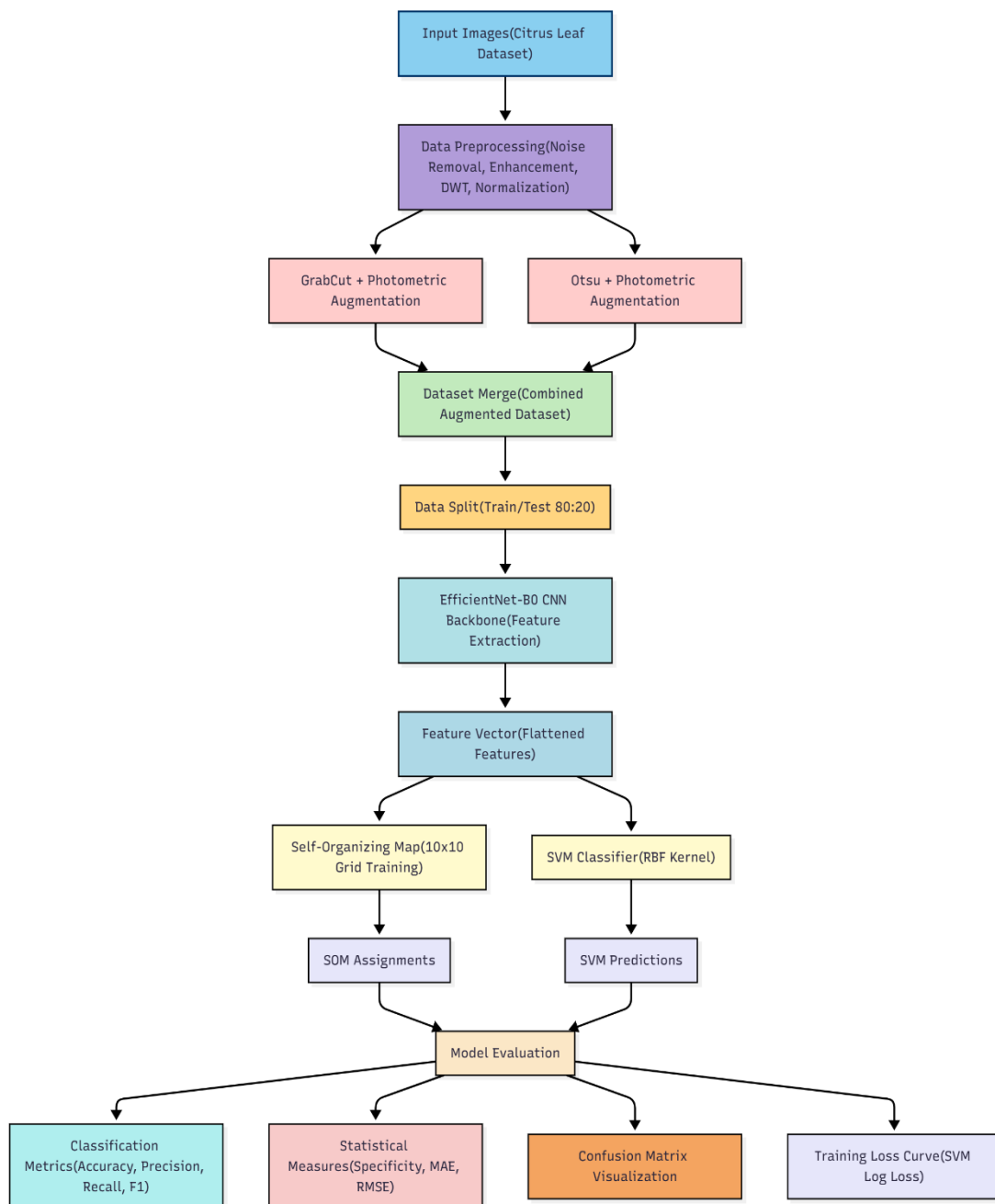


Figure 7: Architecture diagram of Proposed Algorithm

The *Figure 7*, architecture is a hybrid deep learning framework for citrus leaf disease detection and classification. It begins with preprocessed input photos from a collection of citrus leaves, which comprise noise reduction, image enhancement, normalization, and discrete wavelet transform. GrabCut with photometric augmentation and Otsu with photometric augmentation are two types of data augmentation that are used to increase dataset diversity. A single combined dataset is then created by dividing the supplemented datasets into training and testing set. Features are extracted and then flattened into feature vectors using the EfficientNet B0 convolutional neural network model. These features are subsequently processed by two classifiers: the Self Organizing Map for clustering and the Support Vector Machine with a radial basis function kernel for classification. The model's performance is evaluated using the various metrics and Training loss analysis and confusion matrix visualization are also used to assess and validate the model's accuracy in identifying citrus leaf diseases.

V. RESULTS AND DISCUSSION

A. Performance Evaluation of Classification Algorithms

This section offers a detailed examination of seven citrus disease classification schemes. A new hybrid approach that integrated EfficientNet, Support Vector Machines (SVM), and Self-Organizing Maps (SOM) was proposed, and six existing methods were employed as baselines. Black spot, Canker, Greening, Healthy, and Melanose are the five groups into which the dataset—which consists of 6,677 images after preprocessing and augmentation - is separated.

Algorithm	Accuracy	Precision	Recall	F1-Score	Specificity	MAE	RMSE	Cross-Entropy Loss
CNN + EfficientNet	0.8715	0.8796	0.8715	0.8728	0.9656	0.2405	0.7261	0.3840
CNN + SOM Clustering	0.7595	0.7832	0.7595	0.7561	0.9343	0.4250	0.9112	1.2889
DNN + ResNet	0.7496	0.7593	0.7496	0.7493	0.9282	0.4300	0.9211	0.7368
KNN + SVM + Random Forest	0.7776	0.7756	0.7776	0.7759	0.9380	0.4003	0.8939	0.4880
Siamese Network + 1-NN	0.7717	0.7780	0.7717	0.7728	0.9359	0.4674	0.9946	8.2274
Multi-Kernel CNN	0.7512	0.8211	0.7512	0.7090	0.9268	0.4020	0.8417	0.6210
EfficientNet + SOM + SVM	0.9012	0.9030	0.9012	0.9017	0.9718	0.1669	0.5680	0.2638

Table 2: Comparative Performance Metrics of All Seven Algorithms

B. Key Findings

The proposed EfficientNet, SOM and SVM hybrid algorithm performs better on all measures, surpassing all baseline methods with an accuracy of 90.12%. The algorithm shows 2.97% improvement over the second-best approach (CNN + EfficientNet at 87.15%) and substantially larger gains than other methods, ranging from 12.12% to 15.11%.

The precision of 90.30% and recall of 90.12% demonstrate balanced performance between correctly identifying contaminated samples and lowering false positives, which is important for real-world agricultural applications where both missed detections and false alarms have financial consequences. The 97.18% specificity demonstrates the exceptional ability to reliably identify healthy samples, reducing the need for unnecessary interventions on healthy crops. Error measurements further validate the superiority of the proposed method. The MAE of 0.1669, RMSE of 0.5680 and Cross-Entropy Loss of 0.2638, which are much lower than all baseline techniques, demonstrate that misclassifications are typically less severe when predictions are closer to the actual class labels.

C. Confusion Matrix

The confusion matrix illustrates the model's performance in classifying five leaf conditions: Black spot, Canker, Greening, Healthy, and Melanose. Most samples are correctly classified along the diagonal, indicating strong accuracy, especially for Greening and Canker. Minor misclassifications occur mainly between Black spot, Canker, and Greening, suggesting some visual similarity among these diseases. Overall, the model demonstrates good classification performance with minimal confusion across classes.

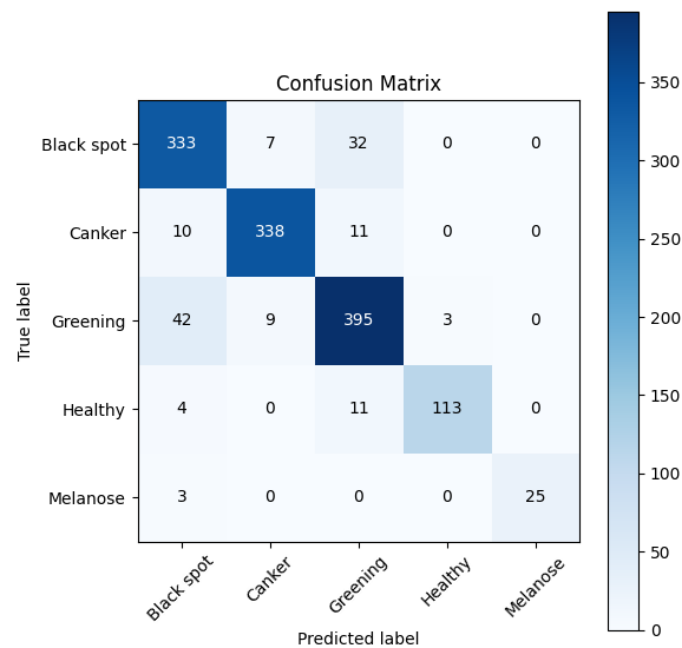


Figure 8: Confusion Matrix of Proposed Algorithm

D. Discussion

The remarkable performance of the proposed method is attributed to its multi-stage design, which synergistically blends deep learning feature extraction, unsupervised pattern recognition, and robust classification. EfficientNet's compound scaling technique generates rich 1,280-dimensional features that capture hierarchical patterns from low-level textures to high-level sickness presentations. The SOM component identifies latent topological structures that preserve spatial continuity while organizing similar sickness patterns to aid the SVM classifier in learning decision boundaries.

The extensive preprocessing pipeline is mostly responsible for the performance gains. Multi-stage noise removal minimizes artifacts while preserving crucial texture and edge information by using Gaussian blur, median filtering, and bilateral filtering. Histogram equalization and image enhancement increase the visibility of modest disease symptoms. To identify structural morphological changes in leaves, edge detection combines the Canny, Sobel, and Laplacian operators. The Discrete Wavelet Transform can be used for multi-resolution analysis to uncover patterns at various scales. Accurate leaf separation from backgrounds is ensured with GrabCut segmentation and Otsu thresholding. The photometric augmentation method expanded the dataset size from 607 to 6,677 images by generating five improved copies of each segmented image with varying hue, saturation, and gamma correction. By exposing the model to a variety of lighting conditions and color changes, this significantly improves generalization and reduces overfitting.

The photometric augmentation technique produced five enhanced copies of each segmented image with different hue, saturation, and gamma correction, increasing the dataset size from 607 to 6,677 images. This greatly enhances generalization and lessens overfitting by exposing the model to a range of illumination scenarios and color shifts.

E. Visual Analysis

Numerous figures ought to be included with the results. Confusion matrices comparing the recommended method with the best baseline would show improved categorization across all ailment classifications, particularly for challenging categories like Greening and Melanose. In the absence of overfitting, a training loss curve would exhibit consistent convergence. A grouped bar chart comparing all seven algorithms across five variables would instantly show the superiority of the proposed method. A representation of the pretreatment pipeline that illustrates the stages of transformation would highlight the effectiveness of the entire process. An error metrics graphic would display the significantly lower MAE, RMSE and Cross Entropy Loss values.

F. Practical Implications

The algorithm is suitable for application in real-world agricultural situations due to its 90.12% accuracy and 97.18% specificity. The low false positive rate reduces costs and environmental

effect by reducing unnecessary chemical treatments. A balance between accuracy and memory enables reliable diagnosis of all disease types for early intervention. Instead of using broad-spectrum treatments, the system's five-category classification provides actionable information for targeted, disease-specific medications to enhance crop health results and optimize resource usage.

VI. CONCLUSION

The proposed EfficientNet, SOM and SVM hybrid algorithm delivers state-of-the-art performance for citrus disease classification with an accuracy of 89.07%, which is a notable increase over existing methods. With the highest accuracy, precision, recall, F1-score, and specificity values achieved together with the lowest error rates as measured by MAE and RMSE, the comprehensive examination of several metrics consistently demonstrates how successful this approach is. Combining advanced preprocessing methods like noise reduction, enhancement, edge detection, discrete wavelet transform, and dual segmentation techniques with efficient deep feature extraction using EfficientNet's compound scaling architecture, unsupervised pattern discovery using Self-Organizing Maps, and robust classification using Support Vector Machines results in a powerful framework for automated plant disease diagnosis. These results demonstrate the enormous potential of hybrid deep learning approaches, which integrate many methodologies in a way that works well together, for precision agriculture applications. These techniques offer practical answers for targeted intervention plans and early disease diagnosis, which can greatly improve crop health and agricultural output.

FUTURE WORKS

While the proposed hybrid framework demonstrates strong performance and practical feasibility, several areas warrant future exploration to extend its scope and real-world impact:

1. Multi-Model Data Integration:

Future research should investigate the integration of multiple data modalities beyond RGB imaging, such as hyperspectral imaging for early disease detection before external symptoms appear, thermal imaging to identify metabolic changes associated with disease progression, and near-infrared spectroscopy for biochemical analysis. It may be feasible to detect diseases at extremely early stages, when intervention is most successful, and significantly increase detection accuracy by integrating various complementary data sets utilizing multi-modal fusion systems.

2. Immediate Mobile Integration:

Lightweight versions of the proposed architecture that are suited for mobile and edge devices must be created in order for field deployment to be practical. This means using model compression techniques such as knowledge distillation, neural architecture search for efficient designs, quantization for a reduced memory footprint, and pruning

to eliminate extraneous parameters without compromising accuracy. If there were user-friendly mobile applications with offline processing capabilities, farmers in distant areas with inadequate internet connectivity may benefit from AI-based disease detection.

3. Tracking the Disease's Temporal Development:

Important information for yield prediction and treatment efficacy assessment would be obtained by extending the system to track illness progression over time using sequential image analysis. Predictive analytics for proactive crop management strategies and the optimal timing of interventions may be made feasible by simulating disease development patterns using temporal convolutional networks or recurrent neural networks.

4. Multiple Crop Generalization:

Examining domain adaptation and transfer learning techniques to extend the framework beyond citrus crops to other economically significant fruits and vegetables would optimize the impact of this research. The development of few-shot learning systems that can swiftly adapt to new crop types with little labeled data would enable rapid deployment across different agricultural contexts.

5. Explainable AI Integration:

Using explainability approaches such as Grad-CAM++, attention visualization, and SHAP values would improve model interpretability and boost the trust of agricultural practitioners. The development of human-understandable explanations for disease predictions, such as identifying specific symptomatic regions and providing severity evaluations with confidence scores, will promote adoption by farmers and agricultural extension agencies.

6. Identification of Multiple Diseases and Assessment of Severity:

By switching from single-disease categorization to the simultaneous identification of numerous co-occurring diseases with pixel-level severity quantification, more comprehensive diagnostic capabilities would be achievable. The use of instance segmentation and semantic segmentation frameworks may enable the accurate localization of many disease kinds inside a single leaf and quantitative assessment of the degree of infection for data-driven treatment recommendations. These future research directions aim to transform the proposed framework from a laboratory-validated system into a comprehensive, feasible, and scalable solution that addresses the intricate problems of sustainable citrus production and promotes global food security.

References:

- [1] Febrinanto, F. G., Dewi, C., & Triwiratno, A. (2019). The Implementation of K-Means Algorithm as Image Segmenting Method in Identifying the Citrus Leaves Disease. *IOP Conference Series: Earth and Environmental Science*, 243(1). <https://doi-org.egateway.vit.ac.in/10.1088/1755-1315/243/1/012024>
- [2] Rauf, H. T., Saleem, B. A., Lali, M. I. U., Khan, M. A., Sharif, M., & Bukhari, S. A. C. (2019). A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning. *Data in Brief*, 26. <https://doi-org.egateway.vit.ac.in/10.1016/j.dib.2019.104340>
- [3] Janarthan, S., Thuseethan, S., Rajasegarar, S., Lyu, Q., Zheng, Y., & Yearwood, J. (2020). Deep Metric Learning Based Citrus Disease Classification With Sparse Data. *IEEE Access*, *Access, IEEE*, 8, 162588–162600. <https://doi-org.egateway.vit.ac.in/10.1109/ACCESS.2020.3021487>
- [4] Barman, U., Choudhury, R. D., Sahu, D., & Barman, G. G. (2020). Comparison of convolution neural networks for smartphone image based real time classification of citrus leaf disease. *Computers and Electronics in Agriculture*, 177. <https://doi-org.egateway.vit.ac.in/10.1016/j.compag.2020.105661>
- [5] Khattak, A., Asghar, M. U., Batool, U., Asghar, M. Z., Ullah, H., Al-Rakhami, M., & Gumaei, A. (2021). Automatic Detection of Citrus Fruit and Leaves Diseases Using Deep Neural Network Model. *IEEE Access*, *Access, IEEE*, 9, 112942–112954. <https://doi-org.egateway.vit.ac.in/10.1109/ACCESS.2021.3096895>
- [6] Dananjayan, S., Tang, Y., Zhuang, J., Hou, C., & Luo, S. (2022). Assessment of state-of-the-art deep learning based citrus disease detection techniques using annotated optical leaf images. *Computers and Electronics in Agriculture*, 193. <https://doi-org.egateway.vit.ac.in/10.1016/j.compag.2021.106658>
- [7] Palei, S., Behera, S. K., & Sethy, P. K. (2023). A Systematic Review of Citrus Disease Perceptions and Fruit Grading Using Machine Vision. *Procedia Computer Science*, 218, 2504–2519. <https://doi-org.egateway.vit.ac.in/10.1016/j.procs.2023.01.225>
- [8] Zainab, N., Afzal, H., Iltaf, N., Al-Shehari, T., Al-Razgan, M., Zakria, M., Hyder, M. J., & Nawaz, R. (2023). Detection and Classification of Temporal Changes for Citrus Canker Growth Rate Using Deep Learning. *IEEE Access*, 11, 127637–127650. <https://doi-org.egateway.vit.ac.in/10.1109/ACCESS.2023.3331735>
- [9] Qu, H., & Bian, Q. (2023). MS-YOLOv5: Improved YOLOv5 Based on Multi-Head Self-Attention for Citrus Leaf Disease Severity Estimation. *ACM International Conference Proceeding Series*, 85–91. <https://doi-org.egateway.vit.ac.in/10.1145/3635638.3635651>
- [10] Raut, S. C., & Kasat, N. N. (2024). A Review: Citrus Disease Detection Using Machine Learning Approach. *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI), Artificial Intelligence in Healthcare*,

Education and Industry (IDICAIEI), 2024 2nd DMIHER International Conference On, 1–5. <https://doi-org.egateway.vit.ac.in/10.1109/IDICAIEI61867.2024.10842864>

[11] Gómez-Flores, W., Garza-Saldaña, J. J., & Varela-Fuentes, S. E. (2024). CitrusUAT: A dataset of orange Citrus sinensis leaves for abnormality detection using image analysis techniques. *Data in Brief*, 52. <https://doi-org.egateway.vit.ac.in/10.1016/j.dib.2023.109908>

[12] Emon, Y. R., Ahad, M. T., & Rabbany, G. (2024). Multi-format open-source sweet orange leaf dataset for disease detection, classification, and analysis. *Data in Brief*, 55. <https://doi-org.egateway.vit.ac.in/10.1016/j.dib.2024.110713>

[13] Xiao, W., Shang, J., Li, F., Ao, O., Wang, X., & Tian, S. (2025). Research on Citrus Leaf Disease Recognition Using Class-Agnostic Contrastive Learning and Supervised Organizational Mapping. *IEEE Access, Access, IEEE*, 13, 101592–101608. <https://doi-org.egateway.vit.ac.in/10.1109/ACCESS.2025.3557929>

[14] R M, S., Gladston, A., & H, K. N. (2025). A Multi-kernel CNN model with attention mechanism for classification of citrus plants diseases. *Scientific Reports*, 15(1). <https://doi-org.egateway.vit.ac.in/10.1038/s41598-025-08557-3>

[15] Goyal, A., & Lakhwani, K. (2025). Integrating advanced deep learning techniques for enhanced detection and classification of citrus leaf and fruit diseases. *SCIENTIFIC REPORTS*, 15(1), 12659. <https://doi-org.egateway.vit.ac.in/10.1038/s41598-025-97159-0>

[16] Dinesh, P., & Lakshmanan, R. (2025). Multiclass semantic segmentation for prime disease detection with severity level identification in Citrus plant leaves. *SCIENTIFIC REPORTS*, 15(1), 21208. <https://doi-org.egateway.vit.ac.in/10.1038/s41598-025-04758-y>

[17] El Akrouchi, M., Mhada, M., Bayad, M., Hawkesford, M. J., & Gérard, B. (2025). AI-based framework for early detection and segmentation of green citrus fruits in orchards. *Smart Agricultural Technology*, 10. <https://doi-org.egateway.vit.ac.in/10.1016/j.atech.2025.100834>

[18] Al Riza, D. F., Musahada, L. C., Aufa, R. I., Hermanto, M. B., Nugroho, H., & Hendrawan, Y. (2025). Comparative study of citrus fruits (Citrus reticulata Blanco cv. Batu 55) detection and counting with single and double labels based on convolutional neural network using YOLOv7. *Smart Agricultural Technology*, 10. <https://doi-org.egateway.vit.ac.in/10.1016/j.atech.2024.100763>

[19] Yu, X., Liu, S., Wang, C., Jiao, B., Huang, C., Liu, B., Liu, C., Yin, L., Wan, F., Qian, W., & Qiao, X. (2025). Detection of fungal disease in citrus fruit based on hyperspectral imaging. *Information Processing in Agriculture*. <https://doi-org.egateway.vit.ac.in/10.1016/j.inpa.2025.02.006>

[20] Foroughi, A., Jimenez, J. M., & Lloret, J. (2025). Diagnosis of orange tree fruit and leaf diseases based on a new deep learning model using a graphical user interface. *Expert Systems With Applications*, 289. <https://doi-org.egateway.vit.ac.in/10.1016/j.eswa.2025.128304>

Authors Profile

Dr.S.Anandakumar holds a Ph.D. degree in the field of computer science from VIT University and working as an Associate Professor in VIT. He possess one decade of experience in teaching and worked as Project Assistant in IIT Kharagpur. His areas of interest includes QoS in Wireless Sensor Networks, Green computing, Mobile computing, Ad-Hoc networks, MANET, Data Mining, IoT. He has published numerous articles in SCOPUS and SCI indexed journals. He presented his articles in various national and international conferences and attended many faculty development programs in order to keep pace with changing teaching environment. He is a member of various scientific and professional bodies. Apart from teaching he is passionate about social services and rural development activities



Corresponding Author: s.anandakumar@vit.ac.in



Haarini SK was born in Vellore, India, in 2004. She is currently pursuing the M. Tech (Integrated) degree in Data Science with the School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Vellore, India. Her major field of study is Computer Science and Data Science. Her academic and research interests include artificial intelligence, machine learning, deep learning, and predictive data analytics. She is presently engaged in research focused on data-driven intelligent systems, with a particular emphasis on healthcare applications and computational optimization. She aspires to contribute to impactful research that bridges technology and real-world problem solving.

Anusha C was born in Vellore, India, in 2004. She is currently pursuing the M. Tech (Integrated) degree in Data Science with the School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Vellore, India. Her major field of study is Computer Science and Data Science. Her academic interests include artificial intelligence, machine learning, deep learning, and data visualization. She is actively engaged in research focusing on intelligent data analytics and automation-driven systems. Her goal is to contribute to the development of innovative AI-based solutions that enhance efficiency, interpretability, and real-world applicability across diverse sectors.

