Journal Pre-proof

A Generalized Deep Learning Approach for Cross-Crop Plant Disease Detection Using the Plant Village Dataset

Roopa R, Rajesh Lingam, Santosh Kumar Ravva, Suresh A, Penubaka Balaji and Avanija J DOI: 10.53759/7669/jmc202505126 Reference: JMC202505126

Journal: Journal of Machine and Computing.

Received 10 March 2025 Revised form 06 May 2025 Accepted 10 June 2025



Please cite this article as: Roopa R, Rajesh Lingam, Santosh Kumar Ravva, Suresh A, Penubaka Balaji and Avanija J, "A Generalized Deep Learning Approach for Cross-Crop Plant Disease Detection Using the Plant Village Dataset", Journal of Machine and Computing. (2025). Doi: https://doi.org/10.53759/7669/jmc202505126.

This PDF file contains an article that has undergone certain improvements after acceptance. These enhancements include the addition of a cover page, metadata, and formatting changes aimed at enhancing readability. However, it is important to note that this version is not considered the final authoritative version of the article.

Prior to its official publication, this version will undergo further stages of refinement, such as copyediting, typesetting, and comprehensive review. These processes are implemented to ensure the article's final form is of the highest quality. The purpose of sharing this version is to offer early visibility of the article's content to readers.

Please be aware that throughout the production process, it is possible that errors or discrepancies may be identified, which could impact the content. Additionally, all legal disclaimers applicable to the journal remain in effect.

© 2025 Published by AnaPub Publications.



A Generalized Deep Learning Approach for Cross-Crop Plant Disease Detection Using the Plant Village Dataset

¹Roopa R, ²Rajesh Lingam, ³Santosh Kumar Ravva, ⁴A Suresh, ⁵Penubaka Balaji, ⁶J.Avanija

¹Assistant Professor, Department of CSE (Data Science), Madanapalle Institute of Technology & Science, Madanapalle, Andhra Pradesh, India.

²Senior Software Developer, Adobe Inc, Boston, U.S.A.

³Associate Professor, Department of CSE, Vasavi College of Engineering, Hyderabad.

⁴Associate Professor, Department of Computer Science and Engineering, GRT Institute of Engineering and Technolog

⁵Assistant Professor, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation Deeme

University), Guntur, Andhra Pradesh, India.

⁶Professor, School of Computing, Mohan Babu University (Erstwhile Sree Vidyanikethan Engineering College), Ter

¹roopa509@gmail.com, ²rlingam@adobe.com, ³santosh@staff.vce.2

⁴csesuresh6@gmail.com, ⁵penubakabalaji.cse@gmail.com, ⁶avans75@ anot

*Corresponding Author: roopa509@gmail.com

Abstract - Plant diseases continue to be one of the leading causes of reduced agricultural pr ctivity worldwide, directly threatening food supply chains and the economic stability of farming communities th the global population steadily increasing, the demand for intelligent, scalable, and highly accurate plant letection systems has never been more critical. Deep learning methods have shown promising results in this field umerous conventional models cannot often generalize well across different crop species and unseen dis limitations hinder their practical deployment in dynamic real-world agricultural environments. In s stud opose a robust and generalized deep e comprehensive and diverse Plant Village dataset. learning-based approach for cross-crop plant disease dete sin Our model is built upon a custom-designed Convolution al Nei ark (CNN) architecture that incorporates a small Inception module. Unlike traditional CNNs, which marily f lobal features of a leaf. Our model detects and as on the across diverse crops and adapting to novel conditions. The analyzes localized disease spread patterns, enhancing small Inception module plays a vital role in enabling m scale feature extraction from small disease-affected patches rchitectural refinement allows the model to learn more without adding excessive computational complexity. This discriminative features, resulting in faster convergence and higher classification accuracy. When trained and validated on wed an impressive accuracy of 98.45%, outperforming many traditional the Plant Village dataset, our model ach high precision, recall, and F1-score, confirming its reliability and approaches. Additionally, it demonstrated robustness. By addressing the challen g and poor generalization, common pitfalls in many deep learning or overfitti models, our method provides a sca ve solution for real-time agricultural disease monitoring. This work and e on agriculture by offering a model that is not only accurate but also generally contributes to the growing field of pre efficient and practical for liverse agricultural settings. Ultimately, our research aims to support the development of smart farmi technol ies that ensure healthier crops and contribute to long-term global food security.

Keywords – Font December December, Deep Learning, Convolutional Neural Network (CNN), Inception Module, Cross-Crop Classificant PlantV. ge Dataset, Image-Based Diagnosis, Disease Localization

I. INTRODUCTION

Agriculture place a vital part in all economic systems and is a basic source of food that sustains life. Improving the agricultual in ustry is essential to increasing output and raising quality. This enhancement necessitates establishing the leadenvironment for crops and plants to flourish healthily. Plant degradation is frequently caused by diseases. According a the United Nations Food and Agriculture Organization, chronic illnesses cost the world economy some \$220 billion a year. They may cause serious harm to the crop or perhaps their total annihilation. Plants can be attacked by bacteria, fungi, viruses, and microscopic animals, which can change their natural form and interfere with their essential processes[1]. Plants can be saved when infections are detected early and neutralized. The crops are safeguarded, and losses are prevented to a greater extent, the sooner they are identified. The conventional disease detection methods, which mostly rely on human diagnosis, are time-consuming and inadequate because of a lack of expertise[2]. The diagnosis must be based on a more trustworthy technique because the data collection method and verification frequency are also insufficient. Modern technologies have been provided as an automatic way to identify plant diseases for this aim [3]. Cutting-edge technologies

like sensors, drones, and robots have drastically altered the way farmers manage their crops [4]. Although machine learning has opened up possibilities for studying data, several issues must be resolved.

Plant disease identification using machine learning and deep learning approaches is a quickly developing subject with encouraging outcomes [5]. Yet since deep learning-based techniques rely on automatic feature extraction rather than human feature selection, they have outperformed other machine learning techniques, particularly in the area of image identification [6]. Deep learning-based techniques for the identification and detection of plant diseases have been proposed in a fee research areas. But there are a lot of barriers that keep this technology from being used more effectively. The impossibility of gathering dataset photos for every illness across all leaf kinds is, in fact, one barrier. Furthermore, a few diseases quickly, making it difficult to catch them on leave in time. Furthermore the model is impacted by the traits of the o p and the disease during the learning process [7]. This suggests that characteristics that can be extracted from one dise and crops cannot be applied to other diseases and crops. A dataset comprising pairs of distinct crops and diseas to train the classification model to develop a generalized method for classifying plant ailments. Regret ly, the sn't a dataset like that, and making one is extremely challenging, if not impossible. The globe is home to of plai and crops, and these organisms are susceptible to millions of illnesses. However, no such dataset exists hd it ficult. if not impossible, to create one. There are millions of plants and crops in the world, and liseases may affect these plants and crops. It takes time and effort to acquire data. These problems dig ly affe the p rmance of deep learning-based systems due to the lack of data. Lack of access to sufficient datas he deep Marning system's hinder capacity to generalize patterns learned in plant disease recognition [8].

Even though many deep learning models have been created for the identification of clan iseases, most of them are specialized for a single crop or dataset, which restricts their applicability to other p t species. They are less useful in real-world settings where farmers might grow multiple crops at once becau r crop-specific concentration. Creating a unified or generic deep learning model that can correctly classify illnes any crop kinds is a major challenge. This calls for a robust architecture that can manage a great degree forms, colors, textures, and illness in 1 signs, as well as a diverse and representative dataset. To ais isst study presents a deep learning-based ldre approach that generalizes plant disease detection across and disease kinds to overcome the drawbacks of cro crop-specific models. Our method concentrates on ccisely d disease rather than depending only on the sight ecting h of damaged leaves. A generalized deep learning mo en for its scalability, efficiency, and adaptability through transfer learning. Using the PlantVillage dataset, we aim nable early diagnosis, assist farmers in low-resource settings, and support sustainable agriculture. The following are the 's significant contributions:

- Generalized Disease Detection Moder, we develop a unified deep-learning model capable of accurately identifying plant diseases across multiple croatype quinating the need for separate models for each crop.
- Our strategy places more emphasis on learning disease-specific patterns than existing approaches, which mostly rely on crop-specific visual as. This proves the model's generalizability and practicality.
- Scalable and Easy to A best for Real-World Use, our model uses transfer learning, which means it can quickly learn to detect diseases in new crop using very little additional data. This makes it a practical, low-cost solution that can be used in a fferent forming environments and helps support sustainable agriculture.
- A new version of the plantV lage dataset was constructed by reprocessing existing images into labeled patches (health v) (unhealth). This restructured dataset, consisting of over 1.8 million patches, enables scalable and flyrible training of goveralized models.
- Our opposet nodel was evaluated alongside other state-of-the-art architectures such as EfficientNet-B0, ResNet-50. MulleNet 2, and Vision Transformer (ViT), providing a thorough comparative performance analysis rejurding to tracy, loss, and generalization behavior.

II. LITERATURE REVIEW

ith the world's population growing and natural resources becoming limited, researchers are using various methods to monitor crops and seedlings, especially to detect diseases. Because deep learning has worked well in many areas, many tudies now focus on using it to diagnose plant and crop diseases. However, researchers face several challenges when developing reliable and effective deep learning models for detecting diseases in plant leaves.

However, Chowdhury et al. [9] proposed a CNN-based model for detecting diseases in plant leaves commonly grown in Bangladesh. They trained the model using online datasets, which were mainly collected from other countries, and achieved

an accuracy of 85.31%. Most existing research has similarly focused on plant data from different regions, not locally sourced samples.

A recent study, Baiju et al. [10]'s Slender-CNN detected key diseases in corn, rice, and wheat with 99.81% accuracy, but missed some less common diseases and crops. Its lightweight design suits mobile and drone use, and it has the potential to expand to other crops and real-time field data. Agriculture plays a vital role in any growing economy, but plant disease threaten their sustainability, especially in crops like tomato and potato. Early detection is crucial, and this study, Kalaivan et al. [11], used machine learning (SVM) and deep learning (ResNet) to identify leaf diseases. After preprocessing and augmentation of the datasets, both models are tested. ResNet outperforms SVM with 94% accuracy and is recommendation for real-time use, along with fertilizer suggestions based on disease type. Aboelenin et al. [12] combined a hybrid nodel, CNNs and Vision Transformers for plant leaf disease detection. Tested on apple and corn datasets from PlantViluge, it achieved accuracies of 99.24% and 98%. The model outperforms many existing methods and shows promitties by deragricultural and computer vision applications.

By incorporating attention mechanisms, Duhan et al. [13] RTR_Lite_MobileNet improve upon Mo accuracy on plant disease datasets by up to 99.92%. Because of its efficiency and porta ws fo -latency real-time detection on devices such as the Raspberry Pi. Additionally, the model shows orld agricultural tentia applications by performing well on complex, real-world data. Mango cultivation pla a vital e in food ecurity and the economy of tropical regions, but it is vulnerable to various leaf diseases that impact quality. Early detection is essential for sustainable farming. In recent studies, Varma et al. [14] have applied deep ning models such as VGG19, InceptionV3, ResNet152V2, DenseNet121, and others to detect mango leaf diseases. Ar ione em, InceptionV3 achieved the highest accuracy of 99.87%, demonstrating its effectiveness compared to other n Els.

In other studies, Kaushik et al. [15] proposed a deep-wise separable-ba d a e neural network (DSDNN) for plant egrates Gaussian filtering, Enthalpydisease detection, particularly in potatoes. Using the PlantVillage data odel he ap based graph clustering, and advanced feature extraction technique chieved 99% accuracy, outperforming validated on the PlantVillage dataset, which may previous methods. Despite its success, the model's perfe S limit its effectiveness in real-world scenarios. Early diseases is essential to prevent crop loss. While ection o torn I on Transformer (ViT) models perform better. This study CNNs have been widely used for disease detection, ent y 13, MobileViT, MViTv2) alongside CNNs (VGG, ResNet, Pacal and Ishak et al. [16] used ViT models (MaxVi DenseNet, Xception) with data augmentation and transfer ning.

Tested on PlantVillage and a new CD&S dataset, the models achieved up An ensemble method further improves accurate to 100% accuracy, outperforming previous netho This approach offers a reliable, automated solution for corn disease CNN model with SE blocks and skip connections, achieving 96% diagnosis. Ashurov et al. [17] introdug mwis accuracy for plant disease detection. is study, Bb a and kumar et al. [18] presented a hybrid model using DenseNet201 and SVM to detect diseases in Corn, at, and Rice. It achieves high accuracy up to 99.82% with a lightweight design, making it effective for real-ti detection. Furthermore, Shafik et al. [19] introduced a hybrid IX-CNN model btion lay combining Inception and Xo for plant disease detection. It achieved up to 100% accuracy on multiple datasets, y Dise such as PlantVillage and Tu e, and over 98% on others. The model uses SVM, DT, and RF classifiers and s. Table 1 summarizes several studies on crop disease detection. supports real

While the torks are we have subjficantly contributed to plant disease detection, several vital gaps remain:

- 1. Generalization cross Crops: The inability of current deep learning models to reliably identify diseases in a viety of the species restricts their applicability in a range of agricultural contexts.
 - Data et Limitations: The PlantVillage dataset does not include real-world changes like ambient noise, lighting variar et, and mixed cropping scenarios. The resilience of the model is weakened by this lack of variance.
- 3. Or ditational Restrictions: Deep learning models have high resource requirements, making it difficult to deploy them on mobile and edge devices, limiting farmers' access.

Table 1: Summary of key studies of	on crop disease detection
------------------------------------	---------------------------

Reference	Dataset	Model	Accuracy (%)	Limitations
[9]	Plant-Village	CNN	85.31	The model struggles to detect multiple diseases affecting the same leaf simultaneously.

[10]	Plant-Village	Slender-CNN	99.81	Real-time learning features are not yet implemented, limiting its ability to improve continuously from new data.
[11]	self	ResNet	94.00	The model was only evaluated using preprocessed images, which may not reflec real-world conditions like poor lighting, background noise, or partially damaged leaves.
[12]	PlantVillage	CNN-ResNet	99.24	The model relies on pre-trained CNNs fur- tuned on ImageNet, which may limit its a lity to fully capture specific disease features ique to certain crops
[13]	PlantVillage	RTR_Lite_MobileNet	99.92	Does not explore how the moder performs in real-time field conditions with a rying lighting and backgrounds.
[14]	MangoLeafBD	InceptionV3	99.87	Although the model show high a concy, its performance inceal-one or field conditions is not evaluate
[15]	PlantVillage	DSDNN	99.00	While the hadel proforms well on the given dataset, its effectiveness in diverse agricultural settings with different crops and disease types remains uncertain.
[16]	self	MaxViT	100.00	Vision Transformers require significant computeronal resources, making real-time replacement of low-power agricultural devices contengin
[17]	PlantVillage	Depthwise CNN + SE Residual	96.	Data imbalance affects model performance; it needs better augmentation and improved interpretability.
[18]	PlantVillage	DenseNet201 + SVM	R 2	Scalability remains a challenge, as the model may not efficiently scale for large-scale agricultural deployment across different farming regions.
[19]	Plant Village	XI-CT	100.00	The model focuses on classification but does not provide insights into disease severity
			Methodo	DLOGY

This section on in s the resented methodology for detecting abandoned objects in various weather conditions. The approach utilizes a vobrid metal and DAWN dataset for object detection and a weather enhancement module to handle adverse situations.

3.1 Datase Pescri,

In the work, which the domain of agricultural image analysis and is primarily used for classifying and recognizing plant leaf iseases [324]. It was selected for its diversity, comprehensiveness, high-quality annotations, benchmark status in plant plant is research, accessibility, and strong relevance to real-world agricultural challenges. The PlantVillage dataset comprises 54,305 RGB images, including healthy and diseased leaves. These images span 14 crop species and cover 20 distinct disease types, categorized into 38 classes, 12 healthy leaf classes, and 26 diseased leaf classes. All images were captured under controlled laboratory conditions with a uniform background and standardized to a resolution of 256×256 pixels, ensuring consistency and ease of processing for deep learning models.

For this study, we utilize an enhanced version of the PlantVillage dataset introduced in [21], in which the leaf regions are segmented from the background to better emulate real-world scenarios and improve the model's robustness in field conditions. This preprocessed version reduces background noise and enables the learning algorithms to focus solely on the

relevant leaf features. Figure 1 illustrates comparative examples of the original images with uniform backgrounds and the segmented images with backgrounds removed.



Figure 1: Examples from the PlantVillage dataset: (left) original images with uniferent or tegrore ds and (right) corresponding images with segmented background

3.2 The Refined PlantVillage Dataset

The updated PlantVillage dataset consists of 1,860,316 image patches divided into 4 unhealthy and 846,162 01healthy samples [21]. Table 2 visualization of the disease name with healthy and unk thy patches. The dataset's labeling was guided by the nature and spread of specific plant diseases. Notably, di ch as Huanglongbing (citrus greening), ases leaf blight (Isariopsis leaf spot), and powdery mildew are known to tire surface area of infected leaves uniformly. Based on this characteristic, all patches extracted from lea se diseases were consistently labeled s e iting as unhealthy, as the presence of the infection was assumed are image region. This labeling strategy ten cross th ensured the dataset's integrity and enhanced the train stness by reducing label noise, particularly for s's T s prò disease types that manifest across full-leaf surface careful selected dataset, a generalized plant disease Using the detection model that can reliably differentiate betwee and diseased leaf sections, irrespective of the crop type, may be trained. Table 3 displays the Number of Sample d Patches per Crop.

Table 2:	Visualizatiop	л me disease	name with	healthy	and u	nhealthy	patches

	Disease Name	Unhealthy Patches	Healthy Patches
	Healthy (All cross	0	647614
	Bacterial spot	97506	72307
	Black rot	23749	38096
	Cedar ar le rust	5831	2273
	Cercospera leaf sp //Gray leaf spot	22738	2400
		56887	2911
	tarly by ht	35947	31302
	cca (Blac, measles)	27737	19499
X	Hanglongbing (Citrus greening)	194626	0
	Late Fight	45011	38440
	af mold	14596	7774
	Leaf blight (Isariopsis leaf spot)	43106	0
	Leaf scorch	39752	1310
	Northern leaf blight	40987	5782
	Powdery mildew	119279	0
	Septoria leaf spot	34377	16983
	Spider mites (Two-spotted spider mite)	15110	26823
	Target spot	6086	38714
	Tomato yellow leaf curl virus	155739	0
	Tomato mosaic virus	5209	2175
	Scab	10998	11986
	Total	1014154	846162



Crop Name	Number of Samples	Number of Patches
Apple	3171	97269
Blueberry	1502	38080
Cherry (incl. sour)	959	64287
Corn (maize)	2857	158735
Grape	3639	127516
Peach	2657	67970
Pepper, bell	2475	91126
Potato	352	76541
Raspberry	371	12723
Soybean	5090	190448
Strawberry	1565	57966
Tomato	23060	625250
Orange	5507	194626
Squash	1835	83457

Table 3: Number of Samples and Patches per Crop

3.3 Proposed Method

This section explains the method used to build a general-purpose system for identifying plant diseases from leaf images. Once trained, the system can recognize signs of infection without needing print pairs about the type of plant or the specific disease. A notable feature of this system is its ability to estimate how much of the leaf is affected by the disease. Figure 2 gives a clear overview of the entire process, showing each key step inpole date the system's work.



Figure 2: Overview of the proposed convolutional neural network (CNN) workflow

3.3.1 Generalization of the process of crop disease detection

To promote generalization across diverse crop types, we introduce a novel approach that removes dependence on cropspecific features and focuses exclusively on disease-related visual cues. Rather than training the model using full leaf images, which inherently contain the crop's identifying characteristics, we decompose each image into small, uniform patches. These patches are sufficiently small to eliminate structural clues related to the crop species, allowing the model to concentrate solely on patterns associated with disease symptoms. Each input image, denoted as I, is split into square patches, referred to as S_1 , S_2 ,..., S_n , where each patch has fixed dimensions of 32×32 pixels. This resolution was carefully selected to be compatible with lightweight convolutional neural network (CNN) architectures, enabling efficient training and inference. To accommodate any image dimension, we resize each image to the nearest multiple of 32 pixels along both height H and width W, then divide it into a grid of patches. The total number of patches per image N_s can be calculated as:

$$N_{s} = \left[\frac{H}{P_{h}}\right] \times \left[\frac{W}{P_{w}}\right]$$
$$I = \{S_{1}, S_{2,\dots,s}, S_{N_{s}}\}$$

where, H and W are the height and width of the resized image, and, P_h and P_w are the patch height and width (structure)).

When a 256×256 -pixel image is split into an 8 by 8 grid, for example, the following is produced:

$$N_s = \frac{256}{32} \times \frac{256}{32} = 64 \ patches$$

We applied a filtering mechanism to eliminate irrelevant or empty patches based on the portion of black pixels in each patch. A pixel is black if its Red, Green, and Blue (RGB) values are all zero. The percent we of black pixels in a given patch S is computed as:

$$B_{\%} = \left(\frac{1}{p} \sum_{i=1}^{p} \delta \left(R_{i} = 0 \land G_{i} = 0\right) B_{i} \neq 0\right) \times 100 \right)$$
(4)

(3)

where, P is the whole number of pixels in the patch, $R_{i,G_{i,B_{i}}}$ are the RGB dues of the *i*th pixel and δ (·) is an indicator function producing 1 if the condition is true and 0 otherwise

Patches with a black pixel percentage of 100%, indicating a couplete at once of leaf content, are discarded. Among the remaining patches, we retain only those whose black fixel accentage is less than or equal to that of the original image (empirically determined to be approximately 50%). This function ensures fidelity to the original dataset distribution while removing noise.

The filtered patch set for a given image is the reefined as:

$$\{S_K \in I \mid B_{\%}(S_k) \le B_{threshold}\}$$

$$(5)$$

where, $B_{threshold} \approx 50\%$.

3.3.2 Generalization Acros. Disease Types

hifts the focus from identifying individual diseases to a more foundational task: detecting whether Our propos rates egard is of the disease type. We achieve this by training the model to distinguish between healthy and a leaf is infect infected how classifying the exact disease. Specifically, we aggregate healthy patches across all crop types to feature representation for healthy leaf tissue. In parallel, we gather infected patches from all disease build a rob the distinguishing characteristics of unhealthy leaves. This setup enables the model to generalize disease cate ies to le or previously unseen disease types. To construct this binary classification framework, all patches derived detecti Village dataset were visually labeled by agricultural experts over three months. Each patch was labeled as the

This binary labeling process is formally represented as:

$$\hat{I} = \{S_i \mid S_i \in I \mid \land healthy \} \cup \{S_k \mid S_k \in I \mid \land unhealthy\}$$
(6)

Using this approach, we created a new dataset version, organized into two categories where:

- H: The set of all healthy patches,
- U: The set of all unhealthy patches.

This restructured dataset can be defined as:

$$H = \sum_{c=1}^{C} \sum_{i=1}^{N_c} \{S_i^C | S_i^C \in I_i^C \land healthy\}$$
$$U = \sum_{c=1}^{C} \sum_{i=1}^{N_c} \{S_i^C | S_i^C \in I_i^C \land unhealthy\}$$

where, C is the number of crop-disease classes in the dataset, N_c is the number of images in class c, I_i^c is the <u>i</u>th images in class c, I_i^c images class c, S_i^c is a patch extracted from I_i^c .

3.3.3 Detection and Quantification of Disease Spread

plied t Once the model is trained to distinguish healthy from unhealthy leaf patches, it can be new leaf images. analy Following the same patching approach, each test image is divided into non-over apping \times 32-pixel patches, and irrelevant patches are discarded. The remaining patches are fed to the classifier, w licts whether each patch is healthy or diseased. This patch-wise prediction not only facilitates disease detection but enables the quantification of the disease's spatial extent on the leaf.

ches predicted as unhealthy N_U and To compute the disease prevalence across the entire leaf, we count the nu healthy N_H. The percentage of diseased area $D_{\%}$ is then calculated as:

(9)

 N_H is Healthy patches. where, N_U is The number of unhealthy (infected) patch

This metric provides a direct, interpretable measure of disease everity on each leaf, enabling more informed agricultural assessments and interventions.

3.3.4 CNN Architecture



Figure 3: The architecture of the small Inception model

The convolutional neural network (CNN) for classifying the extracted leaf patches is based on the small Inception model, a lightweight variation of the original GoogLeNet Inception architecture [22]. GoogLeNet's demonstrated conquest in plant disease detection tasks motivated the selection of this architecture. The smaller variant offers improved computational efficiency and is better suited for small input sizes, such as 32×32 -pixel patches.

The small Inception architecture comprises three primary building blocks:

- **Conv Module**: A standard convolutional layer followed by batch normalization and activation.
- **Inception Module**: A simplified version of the original Inception block, which includes multiple parallel convolution operations with different kernel sizes to capture multi-scale features.
- **Downsample Module**: A module that combines convolution and pooling operations to reduce spatial dimenwhile preserving feature richness.

Figure 3 illustrates a schematic overview of the small Inception architecture, illustrating how these modules organized within the network.

IV. EXPERIMENTAL SETUP AND EVALUATION

This section thoroughly evaluates the proposed disease detection method, emphasizing ion capacity and resilience across various crop species and disease kinds while our core model is a volutio neura etwork (CNN), we also incorporate several advanced architectures-namely Efficientnet-b0, I Mobilenetv2, and Vision et-50 Transformer (ViT)-to conduct a comparative performance analysis. This multi-me pproach ensures a thorough assessment of our method's efficacy by benchmarking it against both traditional and cut edge deep learning models. We begin by testing all models on a newly curated dataset to verify their accuracy in hing between healthy and tine diseased leaf regions across different crops. To further the effectiveness of our pros , we compare our results with stateof-the-art methods that used the PlantVillage dataset. The evaluation em nge of standard classification metrics, ng a holistic view of each model's including accuracy, precision, recall, F1-score, loss, and confusion strengths and limitations.

4.1 Experimental Setup

Python and the TensorFlow framework were used to be in the tool model from the ground up. Extensive experimentation with various combinations was conducted to determine the hyperparameter configuration that produced the best classification accuracy while preserving generalization can bility and avoiding overfitting. The optimization algorithm used was Adam, which was applied with its default parameter. Training and evaluation were conducted on Google Colab, leveraging GPU acceleration for improved computational speed and performance. Table 4 presents the hyperparameter configuration. The dataset was partitioned as follows:

- Training Set: Comprising 80% of the state from each class:
 - 676,929 be thy patches
 - 811,323 un althy parene
- Validation & Test Set prises the remaining 20% of each class:
 - 1 0,233 h h thy atches
 - 22,831 unl althy patches

This balanced and compresensive dataset division ensured sufficient representation of both classes across the training and evaluation phases, mis study clopts a deep learning-based approach using a custom Inception-based CNN to identify plant diseases through learning the classification.



Table 4: Hyperparameter configuration

Hyperparameter	Values
Batch size	32
Epochs	300
Learning rate	0.001
Optimizer	Adam

In our study, we employed several standard evaluation metrics to evaluate the performance of our deep-learning models for plant disease detection.

Accuracy: Accuracy represents the proportion of correctly classified patches (healthy and unhealthy) out of the total patches. It provides an overall measure of classification effectiveness:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: To reduce false positives, precision evaluates the proportion of patches that are truly unhealthy compared to those that were projected to be unhealthy.

$$Precision = \frac{TP}{TP + FP}$$

Recall: Recall measures how well the model detects all critical patches to minimize false negatives.

$$\operatorname{Recall} = \frac{TP}{TP + FN}$$

F1-Score: When the class distribution is unbalanced, the harmonic means of accuracy and recall, or **r**-stree, proveal balanced evaluation of both metrics.

$$F1\text{-score} = 2 \times \frac{Precision \times Recall}{Precision + recall}$$

Loss: Loss quantifies the difference between the predicted labels and the actual labels. Loss **categorical crossentropy loss**, which is standard for multi-class classification problems. Lower values indice better performance.

Confusion Matrix: The confusion matrix thoroughly analyzes prediction outcomes asplaying the proportion of patches accurately or inaccurately categorized as healthy or unhealthy. where,

- **TP**: True Positives (unhealthy patches correctly classified)
- TN: True Negatives (healthy patches correctly cl
- FP: False Positives (healthy patches misclassined as unrealt
- FN: False Negatives (unhealthy patches multissified is healthy,

4.3 Performance Analysis

Assessing testing and training accuracy, together with associated loss values, provides important information about how well each model learns and generalizes. A nodel that maintains low and tightly matched loss values while achieving high accuracy on training and testing datasets in cate and generalization without overfitting or underfitting. Figure 4 depicts the training loss and accuracy curve on this study these metrics serve as a foundation to assess the robustness and adaptability of each deep learning and tecture. The applied to patch-based plant disease detection, as displayed in Table 5.

	•			
Model	Train Accuracy	Test Accuracy	Train Loss	Test Loss
CNN (South Teep, a)	0.9910	0.9845	0.0250	0.0400
Efft. 1. Net-B.	0.9624	0.9395	0.0750	0.1100
ReaNet 10	0.9411	0.9023	0.0950	0.1400
bileNa V2	0.9189	0.8782	0.1400	0.1800
Visi Transformer (ViT)	0.9758	0.9420	0.0500	0.1200

 Table 5: Entromatice analysis of testing and training accuracy along with their loss

Another all evaluated models, the CNN-based Small Inception architecture stands out distinctly, achieving the highest test accurate (98.4.%) with minimal loss (0.0400), closely trailing its training performance (99.10%, loss 0.0250). An antimally charized model is characterized by a near parity between training and test performance, demonstrating its calling to generalize effectively over unknown data without learning the training set. Adapted to localized input patterns, is compact architecture works incredibly well for extracting discriminative features from heterogeneous leaf patches. On the other hand, broader or more profound structures have other drawbacks. Vision Transformer (ViT) showed a little performance difference between training and test loss (0.0500 vs. 0.1200), indicating modest overfitting despite its strength in modeling long-range relationships. Its performance reflects high learning capacity but reduced resilience to patch-level noise, likely due to its dependency on extensive pretraining or larger datasets for full effectiveness.

EfficientNet-B0 achieved a respectable balance (train: 96.24%, test: 93.95%) with moderate loss values (0.0750 and 0.1100), showing stable learning yet slightly less confident predictions under variability. ResNet-50, while structurally

deep, underperformed with a notable gap between training and testing accuracy (94.11% vs. 90.23%), suggesting it may not effectively adapt to the fine-grained texture and shape variations present in disease-infected patches. MobileNetV2, optimized for efficiency, exhibited the most evident signs of overfitting. Although its training accuracy reached 91.89%, the drop in test accuracy to 87.82%, paired with the highest loss values (train: 0.1400, test: 0.1800), indicates difficulty in generalizing to new samples.



Figure 4: Training loss and accuracy progression

In essence, the Small Inception CNN surpasses other models in predictive accuracy and xemplifies architectural alignment with the problem's unique demands of compact, patch-level feature extra action from agricultural data. Its better results highlight the value of creating context-specific models rather than just nodifying generic deep networks, particularly in fields with high inter-class similarity and localized patterns.

Table 6: Performance alyst of the models						
Model	Accuracy	Precisi	Recall	F1-Score		
CNN (Small Inception)	845	0.9852	0.9838	0.9862		
EfficientNet-B0	0.9	0.9410	0.9372	0.9375		
ResNet-50	0.902.	0.8945	0.8987	0.8994		
MobileNetV2	0.8782	0.8505	0.8641	0.8768		
Vision Transform (ViT)	0.9420	0.9383	0.9307	0.9398		

In this study, we compared the perfo of five eep learning models CNN with Small Inception modules, Vision nance Transformer (ViT), EfficientNet-B0, a MobileNetV2 using standard classification metrics such as accuracy, Net-56 yn in Table 6. Among these the CNN model incorporating Small Inception precision, recall, and F1-score modules outperformed all ot ing impressive accuracy of 98.45%, precision of 98.52%, recall of 98.38%, and rs, ach an F1-score of 98.62%. Thi performance can be attributed to the architectural design of the Small Inception xceptior module, which allow rocess image features at multiple scales simultaneously. Instead of using a single Inception module combines convolutions in parallel, enabling the model to extract fine details convolutional l size. terns in one pass. Convolutions reduce dimensionality while adding nonlinearity, allowing as well as l textual icated representations more effectively. the model to rn co

In contrast the Vi an Transformer (ViT), which achieved the second-best results with 94.20% accuracy and an F1-score of 92.98%, reason self-attention mechanisms to model global relationships between image patches. While ViT has shown great access of arge-scale datasets, it lacks certain inductive biases such as locality and shift invariance that are naturally present a CN is. This makes ViT less effective when the dataset is relatively small or lacks sufficient diversity, which explain its slightly lower performance in this case. Moreover, ViT requires more data and training time to generalize tell, which can be a limiting factor for smaller or medium-sized real-world datasets like ours.

EfficientNet-B0 showed competitive performance with 93.95% accuracy and 93.75% F1-score. It benefits from a ompound scaling method that optimally balances depth, width, and resolution. However, B0 is the smallest and most lightweight version in the EfficientNet family, and while it's efficient, it may not have had enough representational capacity to capture the subtle and complex features of mango leaf diseases as effectively as our custom CNN. A larger variant such as EfficientNet-B0 might have yielded better results but at the cost of increased training time and resources.

ResNet-50, known for its deep architecture and residual learning capabilities, achieved 90.23% accuracy and an F1-score of 89.94%. Its performance, although respectable, was likely limited by its complexity relative to the dataset. Deep models like ResNet-50 can suffer from overfitting if the dataset size is not sufficiently large, as they may memorize training patterns rather than generalize to new, unseen data. This trade-off between depth and generalization becomes particularly evident in moderately sized image datasets where simpler, more efficient models may perform better.

MobileNetV2 recorded the lowest performance among the five, with 87.82% accuracy and an F1-score of 87.68 Designed for mobile and embedded systems, MobileNetV2 uses depth wise separable convolutions to minimize computation and model size. While this makes it fast and resource-efficient, the downside is a significant reduct learning capacity. This limitation led to underfitting where the model was unable to fully learn the complex patterns eeded for accurate classification of various mango leaf conditions. We chose the CNN with Small Inception modules for veral important reasons. First, it offers a balanced trade-off between accuracy and computational efficiency. The not as deep or parameter heavy as ResNet or ViT making it easier to train on a modest dataset without the r tting of ov Second, the multi-scale feature extraction capability of the Inception module is particularly effect plant ease classification tasks, where symptoms vary in size, shape, and color. Some disease spots are very l, wh cover d, the reduction in larger leaf areas making a model that can detect both local and global patterns extremely e parameters through 1×1 convolutions before applying larger filters allows the del to htweight while main maintaining high performance, which is ideal for scalable, real-world applications s e disease detection tools. as mo



Figure 6: Performance measurement of the models



Figure 7: Confusion matrix of the presented mode

The confusion matrix in Figure 7 offers a detailed visualization of the predictive performa of the proposed CNN-based Small Inception model on the test dataset, which was balanced across healthy and y leaf patches. The matrix heal reveals powerful classification capability, with True Positives (TP) and True atives (TN) dominating the grid. Specifically, a substantial proportion of unhealthy patches were correctly d as unhealthy (TP), and likewise, the model demonstrated high fidelity in detecting healthy patches (TN). The lse Positives (FP)—healthy samples our misclassified as unhealthy-and False Negatives (FN)-unhealthy s as healthy—remains minimal. These accurate classifications highlight the model's resilience and mo ninor tic errors. The confusion matrix also demonstrates the model's accuracy in detecting the prese hich is essential in agricultural diagnostics because les false negatives can cause severe crop loss and unnoti diseas pread

The confusion matrix confirms that the proposed CNN to be snear-optimal classification across categories, reflecting its strong generalization and discriminative power when apple to real-world plant pathology data.

4.4 Comparative Analysis

To thoroughly evaluate the effectiveness of the appropriate operation of CNN-based Small Inception model, we conducted a comparative analysis against several state-of-the-expansion the widely used PlantVillage dataset presented in Table 7.

Dataset	Mlel	Accuracy	Precision	Recall	F1- score
	eafDise_eNet[20]	97.68	97.50	97.20	97.30
	ption[21]	98.65	98.27	98.04	98.63
Plant ge	CNN+MLP[22]	95.06	95.00	95.00	95.00
\mathbf{V}	CBSNet[23]	92.04	91.58	90.24	90.71
	Proposed Model	98.45	98.52	98.38	98.62

1e 7. Compare our proposed model with existing works

LeafDiseas [20] reported an accuracy of 97.68%, supported by balanced precision (97.50%) and recall (97.20%). performs well, its design primarily focuses on disease classification with relatively limited adaptability Wh his mo r fine-grained feature variations, which are crucial for early-stage disease identification. MXception [21], to pate of the Xception architecture adapted for agricultural applications, achieved a higher accuracy of 98.65%. Its tens separable convolutions improve parameter efficiency, contributing to a strong F1-score of 98.63%. However, espite its remarkable results, MXception's complexity and computational requirements make it less suitable for real-time or embedded applications, especially in resource-constrained agricultural environments. Though simple and computationally light, the CNN+MLP hybrid model [22] lagged with an accuracy of 95.06%. Its low capacity to generalize across many crop-disease scenarios was caused by its flat design, which lacks deeper hierarchical feature extraction. Similarly, CBSNet [23] fared poorly while being designed for crop health monitoring, attaining an accuracy of only 92.04%. Its relatively low F1-score (90.71%) and recall (90.24%) suggest difficulties in reliably detecting sick patches, which an excessive dependence on handmade characteristics and inadequate spatial encoding might cause.

Our proposed model demonstrates a compelling balance between accuracy (98.45%) and computational efficiency. Leveraging a Small Inception architecture effectively captures localized and abstract features within patch-level leaf images. With a precision of 98.52% and a recall of 98.38%, the model maintains exceptional consistency across all metrics, resulting in a robust F1 score of 98.62%, as illustrated in Figure 8. In contrast to deeper networks like ResNet or MXception, our architecture was developed with optimal generalization and inference speed in mind, guaranteeing excellent performance without needless parameter overhead or overfitting.



This study introduces a robust, generalized approach i It disease detection across multiple crop types, grounded in age dataset. Traditional disease classification methods often constructing an updated version of the widely used Plant suffer from limited generalizability, as they tend to be crop-sp fic or reliant on large quantities of disease-specific labeled data. To address this limitation, we reformula he detection problem by shifting focus from identifying individual disease types to distinguishing between healthy an unb leaf regions-regardless of the crop or pathogen. We selected the Small Inception CNN architecture due lance between computational efficiency and high representational oven power, particularly in scenarios inv ng sma ch-based inputs. Its modular structure effectively captures localized oss various crops. This design choice was guided by the need for a scalable and features of healthy and diseased tissue generalizable model to perf across heterogeneous agricultural data. The proposed model demonstrated ate outstanding capability in le disease and cross-crop representations by systematically curating a new binaryning cro labeled dataset from the Plan pus and training on visually annotated leaf patches. This foundation enables more illage (inclusive and ionitoring systems less dependent on specific disease annotations.

Future directions include refining the detection pipeline to incorporate multi-stage learning for disease identification after infection detection, excluding the framework to accommodate real-world, noisy field data, and deploying the model in edge-based environments for real-time, in-field diagnostics. This generalized approach lays the groundwork for more adaptable and intendent agricultural systems, supporting farmers and agronomists in disease prevention and yield prototion on a blobal scale.

REFERENCES

- K. Dutta, D. Talukdar, and S. S. Bora, "Segmentation of unhealthy leaves in cruciferous crops for early disease detection using vegetative indices and Otsu thresholding of aerial images," *Measurement*, vol. 189, p. 110478, Feb. 2022, doi: 10.1016/j.measurement.2021.110478.
- [2] L. Li, S. Zhang, and B. Wang, "Plant Disease Detection and Classification by Deep Learning—A Review," IEEE Access, vol. 9, pp. 56683–56698, 2021, doi: 10.1109/ACCESS.2021.3069646.
- [3] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep Learning for Image-Based Cassava Disease Detection," *Front. Plant Sci.*, vol. 8, p. 1852, Oct. 2017, doi: 10.3389/fpls.2017.01852.

- [4] I. Attri, L. K. Awasthi, T. P. Sharma, and P. Rathee, "A review of deep learning techniques used in agriculture," *Ecological Informatics*, vol. 77, p. 102217, Nov. 2023, doi: 10.1016/j.ecoinf.2023.102217.
- [5] M. Shoaib *et al.*, "An advanced deep learning models-based plant disease detection: A review of recent research," *Front. Plant Sci.*, vol. 14, p. 1158933, Mar. 2023, doi: 10.3389/fpls.2023.1158933.
- [6] B. Liu, Y. Zhang, D. He, and Y. Li, "Identification of Apple Leaf Diseases Based on Deep Convolutional Neur Networks," *Symmetry*, vol. 10, no. 1, p. 11, Dec. 2017, doi: 10.3390/sym10010011.
- [7] Y. Toda and F. Okura, "How Convolutional Neural Networks Diagnose Plant Disease," *Plant Phenomics*, vol. 2019 p. 9237136, 2019, doi: 10.34133/2019/9237136.
- [8] I. Bouacida, B. Farou, L. Djakhdjakha, H. Seridi, and M. Kurulay, "Innovative deep learning approach for cross-cropplant disease detection: A generalized method for identifying unhealthy leaves," *Information Processing in Agriculture*, vol. 12, no. 1, pp. 54–67, Mar. 2025, doi: 10.1016/j.inpa.2024.03.002.
- [9] Department of CSE, Leading University, Sylhet, Bangladesh. *et al.*, "Plant Leaf Disease Detection and Classification Using Deep Learning: A Review and A Proposed System on Bangladesh's Perspective," *IJSB*, vol. 2, no. 1, pp 193– 204, 2023, doi: 10.58970/IJSB.2214.
- [10] B. V. Baiju, N. Kirupanithi, S. Srinivasan, A. Kapoor, S. K. Mathivanan, and M. Astra "Roust CRW crops leaf disease detection and classification in agriculture using hybrid deep learning mucels," *Plat Methors*, vol. 21, no. 1, p. 18, Feb. 2025, doi: 10.1186/s13007-025-01332-5.
- [11] S. Kalaivani, C. Tharini, T. M. S. Viswa, K. Z. F. Sara, and S. T. Abinaya, "Rest Based Classification for Leaf Disease Detection," J. Inst. Eng. India Ser. B, vol. 106, no. 1, pp. 1–14, Feb. 2025, doi:10.1007/s40031-024-01062-7.
- [12] S. Aboelenin, F. A. Elbasheer, M. M. Eltoukhy, W. M. El-Hady, and K. W. Hosny, "A hybrid Framework for plant leaf disease detection and classification using convolutional neural neural neurons are vision transformer," *Complex Intell. Syst.*, vol. 11, no. 2, p. 142, Feb. 2025, doi: 10.1007/s40747-02.01704x.
- [13] S. Duhan, P. Gulia, N. S. Gill, and E. Narwal, "RTR Jin Mode NetV2. Trightweight and efficient model for plant disease detection and classification," *Curra Plan Biology*, vol. 42, p. 100459, Jun. 2025, doi: 10.1016/j.cpb.2025.100459.
- [14] T. Varma, P. Mate, N. A. Azeem, S. Sharma, and Sign, "Automatic mango leaf disease detection using different transfer learning models," *Multimed Tools Appl*, May 024, doi: 10.1007/s11042-024-19265-x.
- [15] I. Kaushik, N. Prakash, and A. Jain, "Plant disease detection using a depth-wise separable-based adaptive deep neural network," *Multimed Tools Appl*, vol. 84(2022, pp. 887–915, Apr. 2024, doi: 10.1007/s11042-024-19047-5.
- [16] I. Pacal and G. Işık, "Utilizing convention of an analysis of the second secon
- [17] A. Y. Ashurov *et al.*, "Enbrying proof disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and restricted sup connections," *Front. Plant Sci.*, vol. 15, p. 1505857, Jan. 2025, doi: 10.3389/fpls.2024.150557.
- [18] A. Bhola and Parly mark Departeature-support vector machine based hybrid model for multi-crop leaf disease identification in Corn, Rice, and Wheat," *Multimed Tools Appl*, vol. 84, no. 8, pp. 4751–4771, Mar. 2024, doi: 10.101/(s/b104, 024-187.5-8.
- [19] W. Shada, A. Tubel, C. Liyanage De Silva, and R. A. Awg Haji Mohd Apong, "A novel hybrid inception-sception constitution: neural network for efficient plant disease classification and detection," *Sci Rep*, vol. 15, no. 1, p. 3936, Jan. 201, doi: 10.1038/s41598-024-82857-y.
- [20] Bouacic , B. Farou, L. Djakhdjakha, H. Seridi, and M. Kurulay, "Innovative deep learning approach for cross-crop plan disease detection: A generalized method for identifying unhealthy leaves," Information Processing in Agriculture, vol. 12, no. 1, pp. 54–67, 2025.
- [21] S. F. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in plant science, vol. 7, p. 215232, 2016.
- [22] M. Dai, W. Sun, L. Wang, M. M. H. Dorjoy, S. Zhang, H. Miao, L. Han, X. Zhang, and M. Wang, "Pepper leaf disease recognition based on enhanced lightweight convolutional neural networks," Frontiers in Plant Science, vol. 14, p. 1230886, 2023.

[23] S. RAJASEKAR, V G., "AN OPTIMIZED ATTENTION-BASED DEEP LEARNING MODEL FOR BLACK GRAM LEAF DISEASE CLASSIFICATION," vol. Vol.103. No.4, Feb. 2025, [Online]. Available: https://www.jatit.org/volumes/Vol103No4/9Vol103No4.pdf

17