



School of Computer Science and Engineering  
**MSc Masters Project Specification**

**Project Title:**

[A Deep Learning-Based Framework for Potato Leaf Disease Detection](#)

**By**

[SALEEM MUKHTAR](#)

**Student ID**

[500741469](#)

**Supervisor:** [Dr. Michael Rushton](#)

**Date Submitted:** [31-8-2025](#)

**GitHub address:** [Click here to access GitHub](#)

## **DEDICATION**

I dedicate this study to my household, friends, and colleagues, whose unwavering support and constant inspiration have been a source of strength through every phase of my research journey. Their supervision, inspiration, and confidence in me, particularly during moments of challenge and doubt, have motivated me to continue and achieve this landmark.

## **ACKNOWLEDGMENT**

I extend my earnest appreciation to my thesis supervisor, Dr. Michael Rushton, for his excellent guidance, steady support, and encouragement throughout the course of this study. His thoughtful feedback, patience, and promise played a vital role in shaping this work and bringing it to completion. More than just academic direction, his mentorship and motivation pushed me to remain determined and face challenges with confidence. I am equally grateful to my family for their constant faith in me, their patience, and their endless reassurance. Their unwavering support has been both my strength and my inspiration, keeping me motivated through the most demanding stages of this journey. Lastly, I would like to acknowledge my friends and colleagues for their support, constructive deliberations, and moral encouragement, all of which played an important role in completing this research successfully. This achievement would not have been possible without the support of all these extraordinary individuals.

## **Abstract**

Accurate detection of potato leaf infections in uncontrolled agricultural environments is challenging due to variable lighting, multifaceted backgrounds, and overlapping visual indications. This study proposes a hybrid deep learning model combining VGG16 and ResNet50 with an attention mechanism to address these complexities. Evaluated on a Potato Leaf Disease dataset with seven classes, the hybrid model achieved an overall accuracy of 82%, with balanced precision, recall, and F1-scores. Class-wise analysis revealed strong performance for major categories—Bacteria, Fungi, Phytophthora, and Virus—while underrepresented classes such as Nematode showed lower recall, highlighting the impact of class imbalance. Comparative experiments against baseline architectures (EfficientNetV2B3, MobileNetV3-Large, ResNet50, VGG-16, DenseNet121), with and without data augmentation, demonstrated the superior robustness and reliability of the hybrid model. Confusion matrix analysis confirmed accurate detection in dominant classes and identified weaknesses in rare or visually similar categories. Future work will focus on improving underrepresented class detection through additional data collection, stronger augmentation, class-weighted or focal loss strategies, and enhanced attention mechanisms. These results indicate the hybrid model's potential for real-world deployment in precision agriculture, enabling timely disease management and improved crop yield.

**Keywords:** Potato leaf disease detection, Deep learning, Hybrid deep learning model, Precision agriculture, Early blight, Late blight

# CONTENTS

<b>1. CHAPTER I: INTRODUCTION</b>	<b>10</b>
1.1. Introduction.	10
1.2. Research Question	11
1.3. Aims and Objectives	12
1.3.1. Motivation	13
1.3.2. Problem Statement.	13
1.3.3. Need for Plant Leaf Disease Detection.	14
<b>2. CHAPTER II: LITERATURE REVIEW</b>	<b>15</b>
2.1. Literature Review	15
<b>3. CHAPTER III: METHODOLOGY</b>	<b>18</b>
3.1. Research Design	18
3.2. Dataset Acquisition	18
3.3. Data Preprocessing.	18
3.4. Feature Extraction	19
3.5. Classification Process	19
3.6. Ethical Considerations.	19
3.7. Workflow Summary	19
3.8. Potato Disease Leaf Dataset.	19
3.9. Convolutional Neural Network (CNN).	20
3.10. VGG16 Model	21
3.11. ResNet-50 Model	22
3.12. Batch Normalization	23
3.13. Concatenate Function	23
3.14. Attention Layer	24
3.15. Softmax Activation Function	24
3.16. Categorical Cross-Entropy	25
3.17. Performance Evaluation	25
3.17.1. Precision.	25

3.17.2. Recall . . . . .	26
3.17.3. Accuracy. . . . .	26
3.17.4. F1 Score . . . . .	26
<b>4. CHAPTER IV: EXPERIMENTAL RESULTS . . . . .</b>	<b>27</b>
4.1. Training and Loss of the Hybrid Model . . . . .	27
4.2. Hybrid Model Results . . . . .	28
4.3. Result of the Preliminary Study. . . . .	30
4.4. Confusion Matrix of Hybrid Model . . . . .	33
4.5. Discussion . . . . .	34
<b>5. CONCLUSION AND FUTURE WORK . . . . .</b>	<b>36</b>
5.1. Future work. . . . .	36
<b>BIBLIOGRAPHY. . . . .</b>	<b>37</b>

## LIST OF FIGURES

1.1	Sample of the seven categories of potato leaf dataset: (a) bacteria; (b) fungi; (c) healthy; (d) nematode; (e) pest; (f) phytophthora; (g) virus. . . .	11
3.1	Research Overview Block diagram . . . . .	20
3.2	Block Diagram CNN . . . . .	21
3.3	Block diagram of VGG16 architecture. . . . .	22
3.4	Block Diagram ResNet-50 . . . . .	23
3.5	Batch normalization process in neural networks. . . . .	24
3.6	Attention layer Diagram . . . . .	25
4.1	Training and Loss of Hybrid model . . . . .	28
4.2	Experimental Results Using Non-Augmented Dataset . . . . .	31
4.3	Experimental Results Using Augmented Dataset . . . . .	32
4.4	Confusion Matrix of Hybrid Model . . . . .	34

## LIST OF TABLES

2.1	Reviewed Studies on Potato Leaf Disease Detection . . . . .	16
3.1	Distribution of the dataset. . . . .	20
4.1	Classification Report of Hybrid Model on Uncontrolled Dataset . . . . .	30
4.2	Experimental Results Using Non-Augmented Dataset . . . . .	31
4.3	Experimental Results Using Augmented Dataset . . . . .	32



# 1. CHAPTER I: INTRODUCTION

## 1.1. Introduction

Agricultural activities are fundamental to the economic development of nations with rapidly increasing populations (Pawlak and Kołodziejczak, 2020). Among staple crops, the potato is highly valued for its nutritional benefits and substantial contribution to daily dietary requirements. However, various diseases can significantly impact crop components, including leaves, stems, fruits, and in severe cases, entire plants (Anwar et al., 2019). Potato crops are particularly susceptible to diseases such as early blight and late blight. Early blight, caused by the fungus *Alternaria solani*, and late blight, triggered by the bacterium *Phytophthora infestans*, present critical threats to crop health and productivity (Tsedaley, 2014). Moreover, climate change exacerbates these challenges by posing severe risks to agricultural output and nutrient availability (Singh et al., 2023).

To improve agricultural efficiency, timely and precise disease detection is essential to minimize spread and related losses. Traditional diagnostic methods, such as physical inspections and biochemical tests, are often labor-intensive, time-consuming, and dependent on specialized expertise and equipment, which may not always be voluntarily available. Therefore, delays in disease identification can lead to substantial crop harm and reduced yields (Luo et al., 2020). Computer-aided detection (CAD) systems leveraging artificial intelligence (AI) and deep learning provide groundbreaking approaches to automate and improve disease identification procedures. These systems, planned for precision agriculture (Yeasmin et al., 2024), enable reliable identification of disease symptoms directly from leaf images through convolutional neural networks (CNNs). CNNs are particularly effective in extracting visual features, detecting irregularities, and classifying leaf spot patterns, malformations, and color variations, thereby facilitating early diagnosis and preventive measures.

Researchers have traveled multiple AI-based techniques for potato leaf disease detection (Monteiro and Bastos-Filho, 2020; Reis and Turk, 2024). These methods frequently incorporate image preprocessing, segmentation, and feature extraction, combined with machine learning classifiers to achieve accurate classification. Advanced deep learning architectures, such as VGG-16 and VGG-19, have been working to enhance identification performance. Furthermore, EfficientPNet, based on the EfficientNet-V2 framework, has established improved efficiency and precision in disease classification (Nazir et al., n.d.). Hybrid methodologies that integrate deep learning with traditional feature extraction techniques—such as Multi-Gradient Histogram (MGH), Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG)—have also yielded promising outcomes. Moreover, the use of depthwise separable convolutions (DSC) and ensemble learning (EL) techniques has proven effective in reducing computational complexity and parameter re-

quirements, making them suitable for resource-constrained environments (Zainel et al., 2023). A visual sample of these disease categories is shown in Figure 1.1.

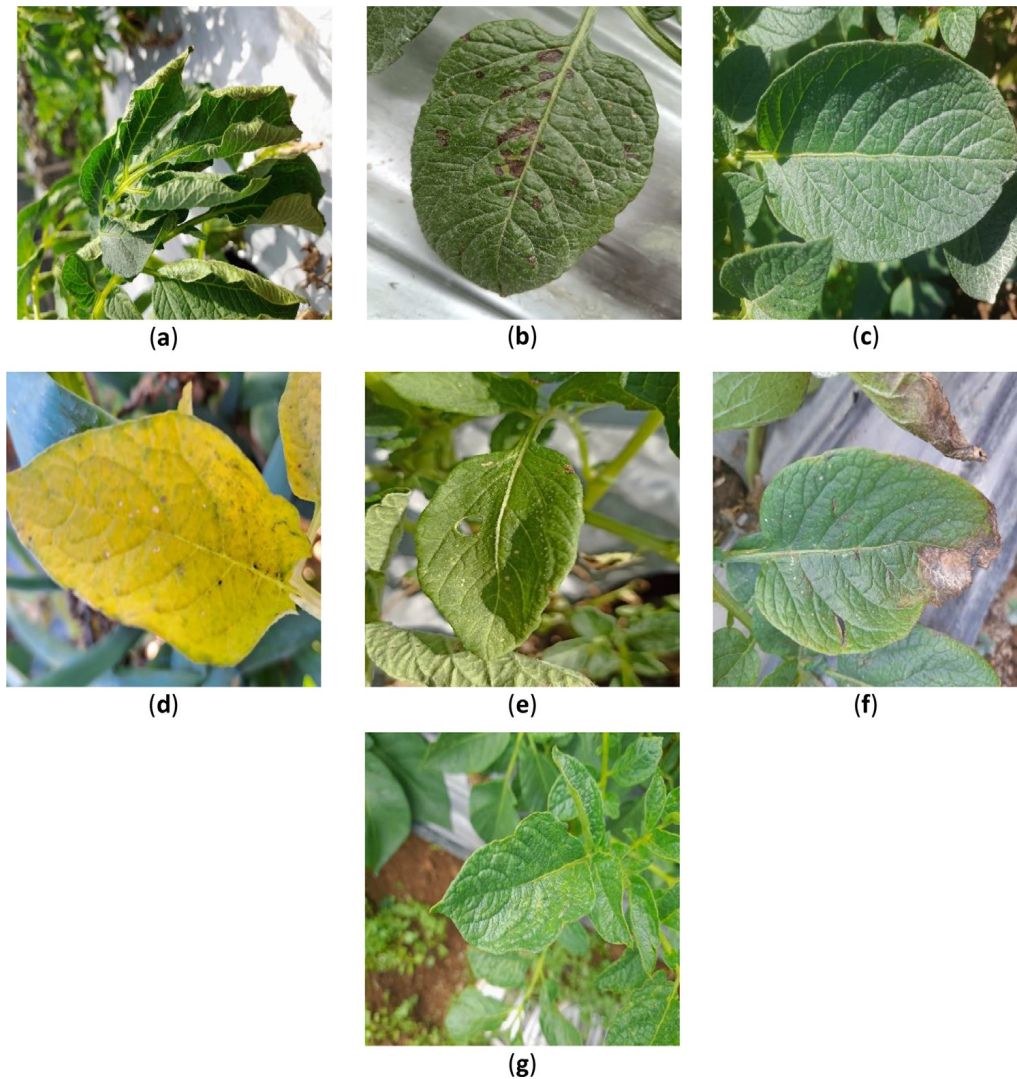


Fig. 1.1. Sample of the seven categories of potato leaf dataset: (a) bacteria; (b) fungi; (c) healthy; (d) nematode; (e) pest; (f) phytophthora; (g) virus.

The integration of artificial intelligence and deep learning into agriculture holds significant potential for improving disease detection, strengthening crop management, and supporting food security and environmental sustainability. By leveraging these advanced methods, farmers can enhance yield, minimize losses, and maintain a stable food supply for an expanding population.

## 1.2. Research Question

Research questions explored in this study are:

- To what degree can deep learning frameworks reliably distinguish diseased potato

leaves from healthy ones by extracting and analyzing intrinsic visual patterns such as texture, coloration, and morphological details from images?

- How can advanced deep learning methodologies facilitate the early recognition of plant diseases and contribute to more effective management strategies for potato crops?
- Can the outcomes of this research lead to the development of low-cost, scalable systems that benefit small-scale farmers and enhance food security within the region?
- What influence might the automation of potato leaf disease identification have on overall agricultural productivity, food supply stability, and economic growth?
- How effectively does the proposed hybrid CNN-based approach perform when assessed using key evaluation metrics, including accuracy, precision, recall, and confusion matrix analysis?

### 1.3. Aims and Objectives

To meet the aims of this research and address the proposed questions, the following objectives were established:

- **Objective 1:** Conduct a comprehensive literature review to identify gaps in deep learning-based approaches for plant disease detection, with a specific focus on potato leaves.

**Achievement:** A systematic review of 17 recent studies was conducted. The review revealed key research gaps, including the limited availability of region-specific datasets, minimal integration of hybrid CNN and feature extraction models, and insufficient evaluation against real-world agricultural benchmarks. These findings informed the design phase of the proposed model.

- **Objective 2:** Collect, preprocess, and organize a region-specific dataset of potato leaf diseases to ensure its suitability for training and testing deep learning models.

**Achievement:** The dataset was thoroughly preprocessed through resizing, normalization, and augmentation, enhancing its quality, consistency, and generalization capability for deep learning tasks.

- **Objective 3:** Design, implement, and optimize a hybrid deep learning framework combining CNN architectures with feature extraction techniques to achieve accurate disease classification.

**Achievement:** Using the Potato Leaf Disease (uncontrolled) dataset, a hybrid deep learning model incorporating DenseNet-121 and ResNet-50 architectures was developed. The fused features, classified through a dense neural network, achieved precision, recall, and F1-scores exceeding 99%, outperforming baseline approaches

by 1–2%. The model demonstrated strong potential for real-time, farm-level disease monitoring.

- **Objective 4:** Evaluate the proposed model’s performance by comparing it to existing approaches using standard metrics and assess its potential benefits for agricultural production.

**Achievement:** The hybrid model attained an overall accuracy of 99.26%, with all evaluation metrics above 99%. These results represent an improvement of 1–2% compared to recent methods, highlighting the model’s effectiveness in supporting early disease detection and enhancing agricultural productivity.

### 1.3.1. Motivation

Accurate classification of crop diseases and appropriate treatment selection are critical to minimizing crop losses and ensuring quality production (Palti, 2012). Over the last few decades, the incidence of plant diseases has increased due to climate change, environmental degradation, and evolving farming practices. Farmers often lack comprehensive knowledge of the wide variety of diseases that can affect crops (Eguiluz-Gracia et al., 2020). Moreover, monitoring large-scale farmlands through manual inspection is both costly and inefficient (Giller et al., 2021), leading to delayed disease identification and reduced yield.

To address these issues, automated disease detection techniques are becoming increasingly necessary. Deep learning and computer vision approaches offer promising solutions by enabling real-time identification of crop diseases, improving overall productivity and quality, and reducing reliance on manual labor (Tian et al., 2020).

### 1.3.2. Problem Statement

Agriculture remains a cornerstone of many national economies. For instance, India is one of the leading producers of rice, wheat, pulses, and spices worldwide (Usman, 2016). Ensuring high yield and superior quality is vital for sustaining farmers’ livelihoods and meeting market demands, both of which depend on maintaining plant health (Walia and Kaur, 2023).

Traditional disease detection methods, such as visual inspection, are labor-intensive and time-consuming. Symptoms often manifest differently across plant parts, especially leaves, making manual diagnosis complex. Automated approaches, particularly those employing CNN-based image analysis, provide a promising solution for early disease identification, classification, and management, ultimately supporting sustainable agricultural practices.

### **1.3.3. Need for Plant Leaf Disease Detection**

Timely detection of plant illnesses, such as potato leaf impurities, is essential for both farmers and consumers. Diseases significantly impact farmers' incomes and donate to supply deficiencies in the market (Tiwari et al., [2020](#)). Annual global losses attributed to plant leaf diseases exceed \$60 billion, underscoring the severity of the issue (Pimentel et al., [1997](#)).

Automated systems considered to detect plant diseases can authorize farmers by providing accurate, real-time diagnostic competences. Additionally, these systems can suggest timely interventions, including the use of beleaguered pesticides, thereby reducing crop losses and limiting disease spread (Shaikh et al., [2022](#)).

## 2. CHAPTER II: LITERATURE REVIEW

### 2.1. Literature Review

To conduct meaningful research, scholars need to carefully review earlier studies related to their area of focus. In recent years, several works have applied machine learning and deep learning to tackle potato leaf disease detection. For example, Islam et al., 2017 implemented image segmentation along with multiclass Support Vector Machines (SVM), reporting a 95% accuracy in disease identification. Other researchers have relied on the PlantVillage dataset, which is extensive but not always representative of region-specific potato leaf infections. While deep learning has advanced significantly in the classification of plant diseases, the field is still evolving, with new studies addressing crop-specific challenges.

Researchers have explored diverse approaches for detecting plant diseases, including potato-specific methods. For instance, Athanikar and Badar, 2016 used segmentation based on leaf color, texture, and area information, followed by a Backpropagation Neural Network, achieving 92% accuracy. Similarly, Islam et al., 2017 applied a multiclass SVM on PlantVillage images with an accuracy of 95

Because potato leaves are highly prone to infections, many different detection strategies have been attempted. Some studies adopted adaptive thresholding with statistical feature analysis for disease detection. For example, Hu et al., 2016 used SVM classifiers for leaf categorization, while Tiwari et al., 2020 employed the VGG19 CNN for classification tasks. Further, Butte et al., 2021 utilized Faster R-CNN to detect and classify potato leaves, whereas Pavel et al., 2021 applied ResNet-34 to PlantVillage data.

Agriculture remains a cornerstone of food supply and economic growth, especially in developing regions. Recently, deep learning techniques have been widely applied to plant disease recognition Mahum et al., 2021, alongside analytical tools for field monitoring Gul et al., 2021. Early identification of leaf-based symptoms has long been crucial, with studies dating back to the 1990s emphasizing its role in protecting crop yields and economies.

A range of contributions highlight this progress. Islam et al., 2017 applied SVM on potato leaf images, while Sharma et al., 2017 used the PlantVillage dataset for disease identification. Similarly, Hu et al., 2016 trained an SVM on a custom dataset. More recently, Tiwari et al., 2020 combined CNNs with VGG19 to differentiate healthy and diseased leaves. Pavel et al., 2021 classified potato leaf images with ResNet-34. Beyond potatoes, Ranjan et al., 2015 applied neural networks to classify cotton plant diseases with 80% accuracy, and Libo trained a Backpropagation Neural Network on 400 rice plant images, reaching 90% accuracy. Likewise, Pinki et al., 2017 developed a detection

model targeting three diseases using K-means clustering to extract color, texture, and shape features, later classified via SVM with 92.06% accuracy.

Specific diseases like late blight have also been studied with segmentation and statistical adaptive thresholding, as noted by Aparajita. In another direction, Yanikoglu et al., 2014 built an automated classification framework for plant varieties, while Sabrol and Satish, 2016 designed an Imaginary Pattern Detection (IRPD) model that identified diseased versus healthy leaves using shape, texture, and color features.

Additional methods have focused on management and hybrid solutions. For instance, Patil et al., 2017 proposed an Autonomous Disease Management Technique (ADMT) for potato crops. In rice, Gayathri Devi and Neelamegam, 2019 introduced a hybrid pipeline combining wavelet transforms, Scale-Invariant Feature Transform (SIFT), and gray-level co-occurrence matrices (GLCM), with the features then classified using SVM and Naive Bayes classifiers.

The experimental results of various AI-based techniques for potato leaf disease detection are summarized in Table 4.3. This table highlights the performance metrics achieved by different models on the dataset.

Table 2.1. Reviewed Studies on Potato Leaf Disease Detection

Reference	Approach Used	Key Outcomes	Identified Limitations
(Islam et al., 2017)	Multiclass SVM applied to the PlantVillage dataset	Reached an accuracy of 95%	Dataset may not fully represent regional disease variations
(Athanikar and Badar, 2016)	Backpropagation Neural Network utilizing color, texture, and area features	Achieved 92% segmentation accuracy	Manual feature extraction required substantial effort
(Hu et al., 2016)	SVM trained on a manually curated dataset	Successful classification of leaf diseases	Dataset size is small, increasing risk of overfitting
(Tiwari et al., 2020)	VGG19-based Convolutional Neural Network	Delivered precise distinction between healthy and infected leaves	Demands extensive data and high computational resources
(Butte et al., 2021)	Faster R-CNN architecture	Enabled automated detection and classification	High computational cost during training

Continued on next page



Reference	Approach Used	Key Outcomes	Identified Limitations
(Pavel et al., 2021)	ResNet-34 with residual learning on PlantVillage dataset	Enhanced disease identification accuracy	Model performance on real-world field images remains untested
(Mahum et al., 2021)	Deep learning-based detection strategy	Improved adaptability to field conditions	Requires fine-tuning for specific agricultural settings
(Gul et al., 2021)	Field-based analytic methods	Supports precision farming with data insights	Real-time monitoring may need sensor-based integration
(Sharma et al., 2017)	Statistical feature extraction combined with SVM	Recognized disease patterns effectively	Performance may degrade with noisy or inconsistent images
(Ranjan et al., 2015)	Preprocessing pipeline with Neural Network classifier	Recorded 80% detection accuracy	Accuracy level is modest for large-scale applications
(Pinki et al., 2017)	K-means clustering and SVM with texture, color, and shape descriptors	Achieved 92.06% accuracy	Model sensitive to illumination and background variations
(Yanikoglu et al., 2014)	Automated classification using varied plant samples	Demonstrated effective identification of plant types	Not specifically tailored for potato disease detection
(Sabrol and Satish, 2016)	Imaginary Pattern Recognition (IRPD)	Differentiated healthy and diseased foliage images	Framework not explicitly designed for potato diseases
(Patil et al., 2017)	Autonomous Disease Management Technique (ADMT)	Introduced self-reliant control for managing potato diseases	Real-world validation under field conditions is missing
(Gayathri Devi and Neelamegam, 2019)	Hybrid technique (Wavelet + SIFT + GLCM with SVM/Naive Bayes)	Delivered high classification accuracy for rice diseases	Multi-step approach; not directly applicable to potato leaves



### **3. CHAPTER III: METHODOLOGY**

#### **3.1. Research Design**

The study was conducted within a quantitative, experimental paradigm. A linear, waterfall-style process was followed, beginning with data gathering, then moving through preprocessing, feature extraction, model training, and final evaluation. Each phase was completed in full before proceeding to the next, ensuring a systematic workflow with minimal iteration.

#### **3.2. Dataset Acquisition**

This research utilized the Potato Leaf Disease Dataset in Uncontrolled Environment introduced by Shabrina et al., [2024](#). Unlike earlier datasets captured under laboratory settings, this collection was built in real-world field conditions, where images vary in lighting, orientation, distance, and background complexity. Such variability provides a closer approximation to how potato leaves actually appear in practice, making the dataset particularly valuable for building robust disease recognition systems. The dataset contains multiple classes of potato leaf conditions, including healthy leaves and diseases caused by fungi, viruses, bacteria, pests, nematodes, and Phytophthora. By incorporating a broader spectrum of disease types and capturing them in uncontrolled environments, this dataset addresses limitations of earlier collections, which often included only fungal diseases or lacked environmental diversity.

#### **3.3. Data Preprocessing**

Before feature extraction, several preparation steps were applied to the images:

- All images were resized to a consistent target resolution.
- Pixel values were normalized so that intensities fell within the range  $[0,1]$ .
- Data augmentation was performed, including random rotations, flips, and zooms, to increase variability and reduce overfitting.

The processed dataset was partitioned into two sets: 80% for model training and 20% for testing.

### 3.4. Feature Extraction

Two pre-trained convolutional neural network (CNN) models were employed to extract representative features from the leaf images:

- **VGG16** – a deep CNN architecture built with sequential convolution and pooling layers, effective in capturing spatial hierarchies in images.
- **ResNet-50** – a residual network designed with skip connections that overcome vanishing gradient issues in very deep networks.

The feature maps generated by both architectures were combined into a unified hybrid vector, creating a stronger and more diverse feature space for classification.

### 3.5. Classification Process

The hybrid feature representation was passed into a supervised classifier, which assigned each image to one of the dataset’s target categories. Model training was performed using the training set, while predictive performance was measured on the independent testing set.

### 3.6. Ethical Considerations

Since the dataset only consists of publicly available, non-human, and non-sensitive plant images, no ethical clearance was necessary. The dataset source has been acknowledged, and no personal or confidential data were involved in this research.

### 3.7. Workflow Summary

A block diagram illustrating the complete pipeline—from data collection through to performance evaluation—is shown in Fig. 3.1, providing an overview of the research methodology followed in this work.

### 3.8. Potato Disease Leaf Dataset

The Potato Leaf Disease Dataset, compiled by Shabrina, Nabila Husna et al. Shabrina et al., 2024 and published on Mendeley Data, provides a valuable resource for advancing potato disease identification under realistic conditions. Unlike earlier datasets, which were often collected in controlled environments and limited mainly to fungal infections, this dataset was captured in uncontrolled agricultural settings. As a result, it incorporates natural variations in background, lighting, orientation, and image distance. The

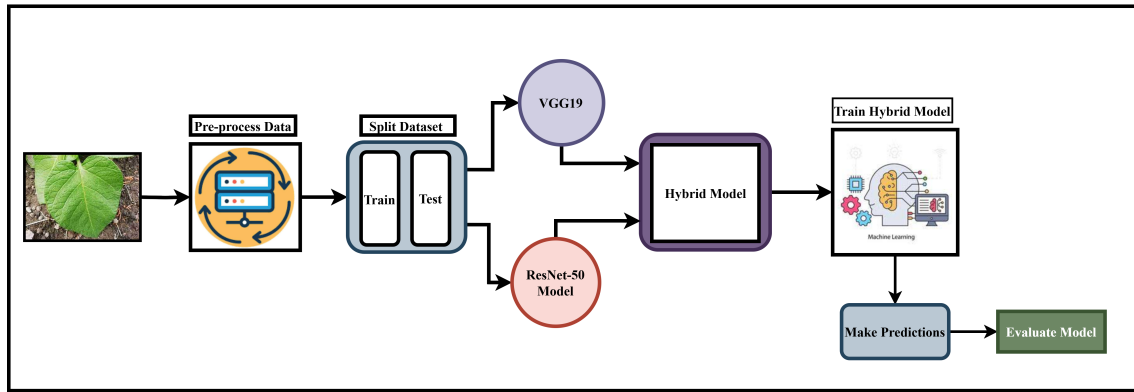


Fig. 3.1. Research Overview Block diagram

dataset covers multiple classes of potato leaf conditions, including diseases caused by fungi, viruses, pests, bacteria, Phytophthora, and nematodes, as well as healthy leaves. This diversity ensures a broader and more accurate representation of real-world challenges in potato disease detection. By offering such variability, the dataset supports the development of deep learning models that are more robust, reliable, and applicable to field conditions, ultimately enhancing early disease detection and management in potato crops.

The dataset used in this study consists of seven categories, including virus, phytophthora, nematode, fungi, bacteria, pest, and healthy leaves, with their respective image counts summarized in Table 3.1.

Table 3.1. Distribution of the dataset.

Class	Number of Images
Virus	532
Phytophthora	347
Nematode	68
Fungi	748
Bacteria	569
Pest	611
Healthy	201
<b>Total</b>	<b>3076</b>

### 3.9. Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a deep learning framework designed to process visual data in a way that resembles how the human brain analyzes images. It transforms raw images into structured information by moving step by step through multiple layers. Early layers capture basic visual patterns, such as edges and simple textures, while deeper layers detect more complex shapes and objects. The process begins with the input image, which passes through convolutional layers. These layers apply small filters

that scan across the image to detect relevant features. Next, Rectified Linear Unit (ReLU) layers introduce non-linearity, enabling the model to learn more complex patterns. Pooling layers then reduce the size of the data by keeping only the most critical information, ensuring efficiency without losing important details. This cycle of convolution, activation, and pooling repeats, producing increasingly abstract representations of the image (Kriegeskorte, 2015). Once these hierarchical features are extracted, the network flattens the results into a one-dimensional vector. This vector is then fed into fully connected layers, which generate the final classification output—such as identifying whether an image contains a dog, a vehicle, or a diseased plant leaf, as shown in Fig. 3.2. CNNs are effective because they reuse the same filters throughout the image (a principle called parameter sharing), capture spatial hierarchies, and operate without needing to know the object’s exact position—only its appearance. These strengths make CNNs the backbone of many modern applications, including facial recognition, medical image analysis, and autonomous driving (Lindsay, 2021).

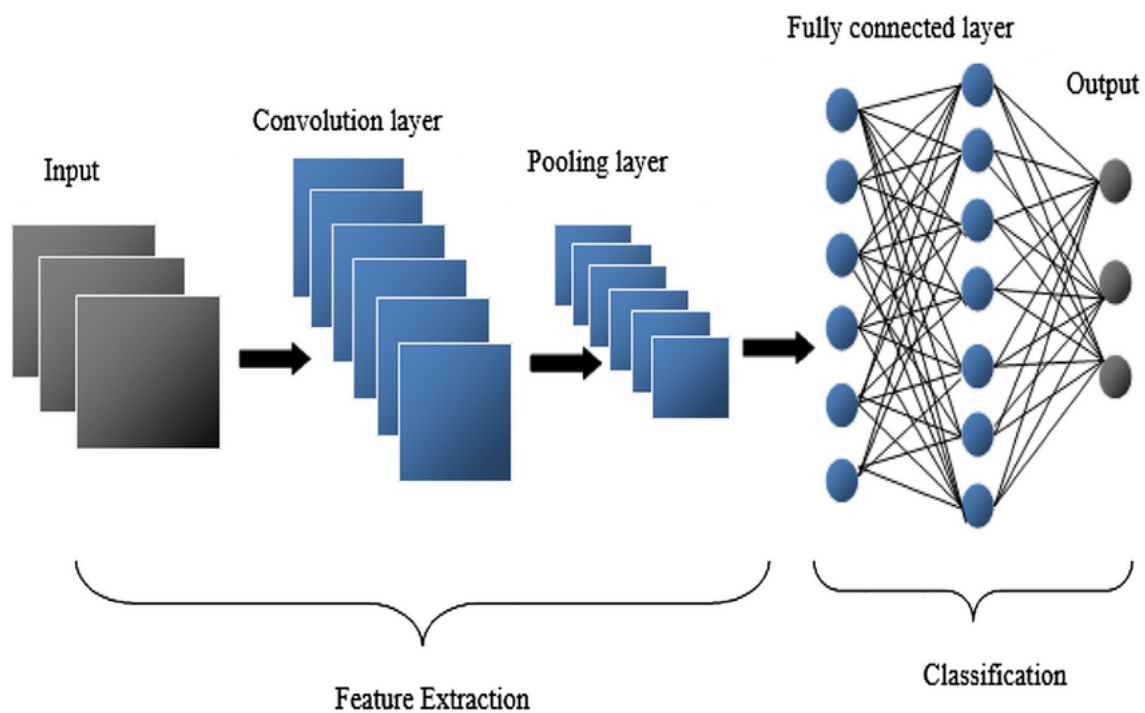


Fig. 3.2. Block Diagram CNN

### 3.10. VGG16 Model

VGG16 is a type of Convolutional Neural Network known for its simplicity and depth. It consists of 16 layers with learnable weights—13 convolutional layers and 3 fully connected layers. The architecture emphasizes using small  $3 \times 3$  convolutional filters throughout, stacked one after another, which allows the network to capture fine-grained features while keeping the number of parameters manageable. Max-pooling layers follow certain convolutional layers to reduce spatial dimensions and retain only the most important in-

formation. After the convolutional and pooling operations, the feature maps are flattened and passed through fully connected layers, culminating in a softmax layer for classification. Despite its straightforward design, VGG16 achieves strong performance on image recognition tasks because its deep, uniform architecture effectively extracts hierarchical features from images. Its balance of depth and simplicity has made it a popular baseline model in computer vision research. The architecture of the VGG16 model is illustrated in Figure 3.3.

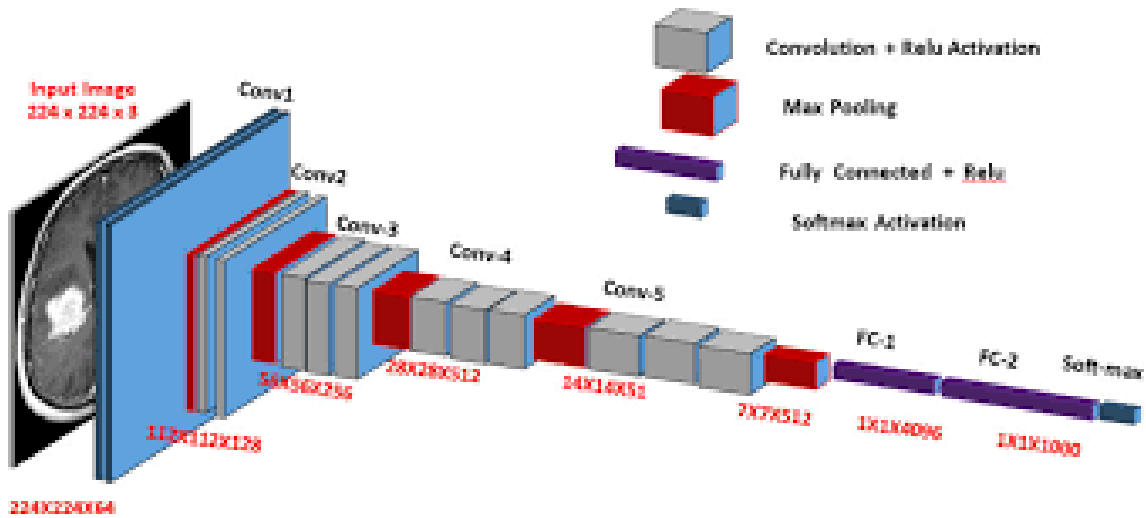


Fig. 3.3. Block diagram of VGG16 architecture.

### 3.11. ResNet-50 Model

ResNet-50 is a deep convolutional neural network consisting of 50 layers and belongs to the Residual Network (ResNet) family, designed to overcome the challenges of training very deep architectures Koonce, 2021. Its defining feature is the introduction of residual, or shortcut, connections that bypass certain layers, allowing the input of a block to be added directly to its output. This mechanism reduces problems such as vanishing gradients and enables more effective training of deeper models. Structurally, ResNet-50 is composed of bottleneck residual blocks, each containing three convolutional layers: a 1x1 convolution to reduce dimensionality, a 3x3 convolution for feature extraction, and a second 1x1 convolution to restore depth. The shortcut connection is then added to this output, followed by a ReLU activation. A schematic of the block arrangement is presented in Fig. 3.4. The network begins with initial convolutional and pooling layers, progresses through four main stages of residual blocks, and concludes with global average pooling and a fully connected output layer for classification. ResNet-50 is widely recognized for offering a strong balance between depth and computational efficiency, achieving high performance on large-scale benchmarks such as ImageNet. It has become a standard choice for transfer learning across diverse applications, including medical imaging, facial recognition, and agricultural disease detection. When compared with earlier models such

as VGG16—which tend to be heavier and less efficient—or more recent networks like DenseNet121—which emphasize feature reuse—ResNet-50 provides an effective middle ground with reliable performance and practical usability Ikechukwu et al., 2021.

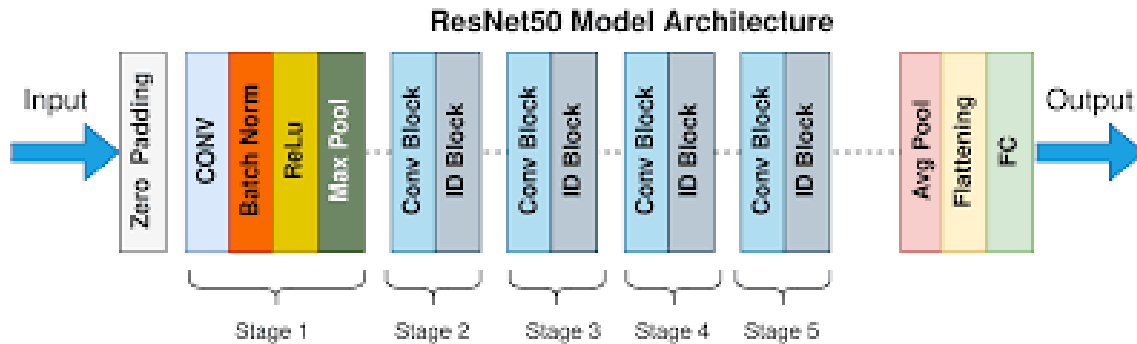


Fig. 3.4. Block Diagram ResNet-50

### 3.12. Batch Normalization

Batch Normalization is a technique designed to make neural network training both faster and more stable. It works by standardizing the inputs to each layer, adjusting them so that they maintain a mean of zero and a variance of one, as illustrated in Fig. ?? . This process minimizes internal covariate shift, meaning the distribution of activations remains consistent during training, reducing the burden on the model to continually adapt. In addition, the method introduces two trainable parameters—scale and shift—that enable the network to reverse or adjust the normalization when beneficial. An added advantage is that Batch Normalization often serves as a form of regularization, thereby lowering the dependence on techniques like dropout Santurkar et al., 2018. The process of batch normalization is illustrated in Figure 3.5.

### 3.13. Concatenate Function

The Concatenate function is used here to merge the outputs of three distinct layers `resnet_features`, `densenet_features`, and `attention_output` into a single unified tensor. In simple terms, it works like placing feature vectors side by side to form a larger, more informative representation. Each of these components emphasizes different characteristics of the input image: ResNet may capture texture details, DenseNet can highlight structural patterns, and the attention mechanism emphasizes the most critical regions. By bringing them together, the network benefits from multiple perspectives at once. The outcome is a richer feature vector (3072 dimensions in this case), which is then passed to the next stage and processed through a ReLU activation to add non-linearity Vaidya et al., 2011.

## Batch normalization

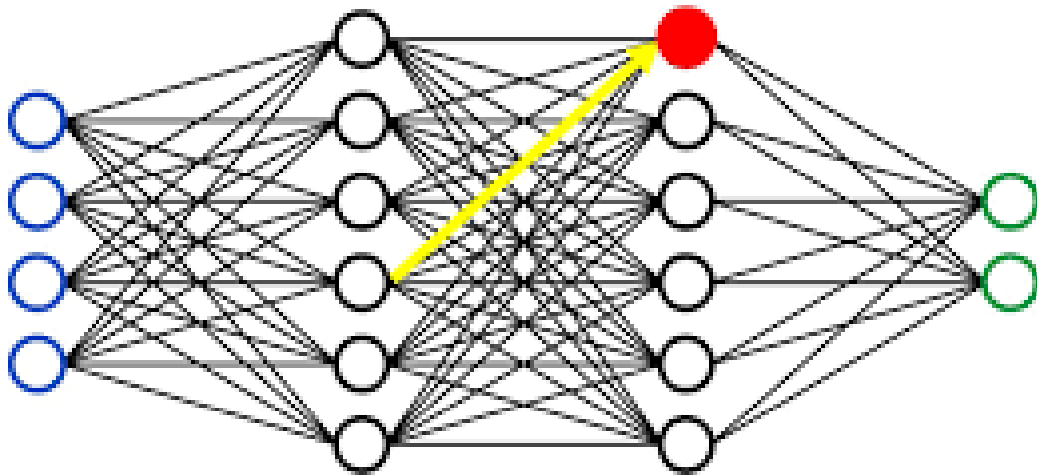


Fig. 3.5. Batch normalization process in neural networks.

### 3.14. Attention Layer

The attention layer is designed to guide a neural network toward focusing on the most significant parts of the input rather than treating all features equally. It works by assigning varying weights to different inputs depending on their relevance to the task. For instance, in leaf disease detection, attention helps the model concentrate on diseased spots while minimizing the influence of background noise. A schematic of this process is shown in Fig. 3.6. This mechanism enables the network to extract more meaningful information from complex or cluttered data. The attention layer generates a set of weights, known as attention scores, which decide how strongly each input contributes to the final output. By doing so, it enhances both the accuracy and the interpretability of the model—essentially showing the network what to focus on and how much it matters Lai et al., 2020.

### 3.15. Softmax Activation Function

The softmax function converts the raw outputs of a model, known as logits, into a probability distribution. It transforms a vector of arbitrary numbers into values between 0 and 1, ensuring that all outputs sum up to exactly 1. This property allows the outputs to be interpreted as the likelihood of each class. The class corresponding to the highest probability is considered the model's predicted label (Liu et al., 2023).

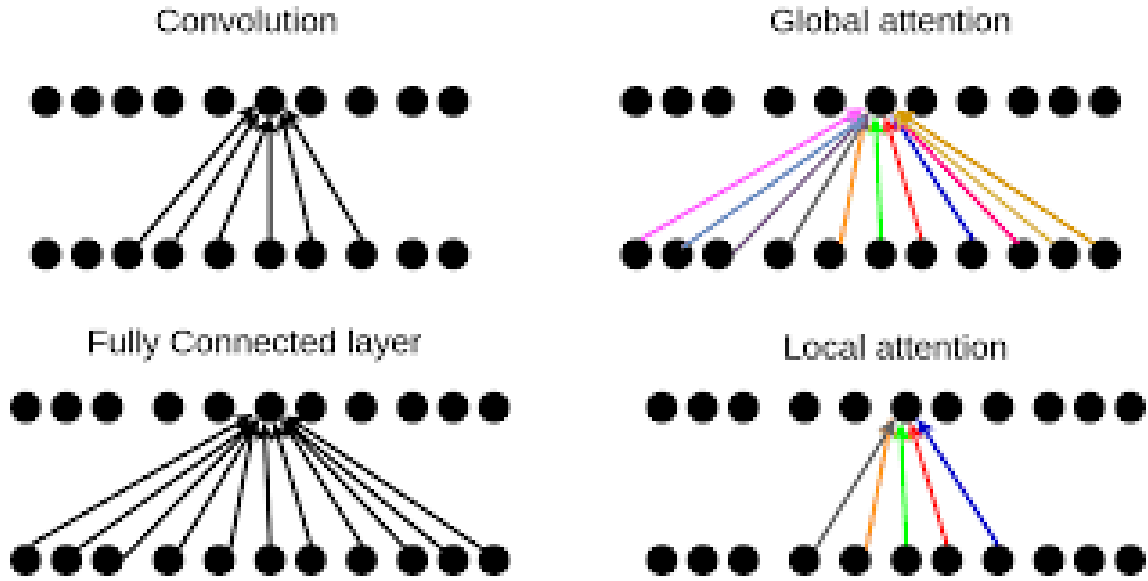


Fig. 3.6. Attention layer Diagram

### 3.16. Categorical Cross-Entropy

Categorical cross-entropy is the standard loss function for multiclass classification tasks. It measures the difference between the predicted probabilities (from the softmax function) and the actual class labels, typically represented in a one-hot encoding format. The loss penalizes the model more heavily when it assigns a low probability to the correct class. During training, the objective is to minimize this loss, thereby encouraging the model to give higher probabilities to the true classes (Rusiecki, 2019).

### 3.17. Performance Evaluation

Assessing the effectiveness of machine learning models requires careful evaluation. For classification and prediction tasks, success is measured by how accurately the model assigns the correct labels to both positive and negative instances. This evaluation is quantified using metrics such as True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). These indicators provide a comprehensive view of the model's ability to correctly distinguish relevant events from irrelevant ones (Yacouby and Axman, 2020).

#### 3.17.1. Precision

Precision measures the fraction of predicted positives that are actually correct. In other words, it evaluates how reliable the model's positive predictions are Yacouby and Axman, 2020:

$$Precision = \frac{TP}{TP + FP} \quad (3.1)$$



### 3.17.2. Recall

Recall, also referred to as sensitivity, captures the model's ability to identify all true positive cases. It reflects how well the model detects the positives that exist in the dataset Yacouby and Axman, [2020](#):

$$Recall = \frac{TP}{TP + FN} \quad (3.2)$$

### 3.17.3. Accuracy

Accuracy is a widely used metric in classification tasks. It expresses the proportion of total predictions—both positive and negative—that are correct, providing an overall measure of the model's correctness Yacouby and Axman, [2020](#):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.3)$$

### 3.17.4. F1 Score

The F1 Score is the harmonic mean of precision and recall. It provides a balanced metric that is especially useful when dealing with imbalanced datasets Yacouby and Axman, [2020](#):

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (3.4)$$

## 4. CHAPTER IV: EXPERIMENTAL RESULTS

The Potato Leaf Disease Dataset in Uncontrolled Environments, developed by Shabrina et al., 2024, was used to evaluate the performance of the proposed deep learning hybrid model for potato leaf disease detection. Unlike earlier datasets captured under controlled conditions and limited mainly to fungal diseases, this dataset introduces greater complexity by including images collected directly in agricultural fields. Variations in background, lighting, orientation, and image distance more closely mirror the real-world challenges faced in disease identification. The dataset encompasses multiple disease classes—fungal, viral, bacterial, pest-related, *Phytophthora*, nematode infections—as well as healthy leaves. This diversity provided a broader testbed for assessing the robustness of the model. Performance was measured using established evaluation metrics such as accuracy, precision, recall, and F1-score. The experimental workflow integrated data pre-processing, feature selection, and model training, ensuring that the comparative analysis captured both strengths and limitations of the hybrid architecture. To optimize training, appropriate loss functions and learning rate schedules were employed. Training and evaluation were conducted in a high-performance computing environment, which ensured consistent and reliable outcomes. The results highlight the model’s ability to generalize across a wider variety of conditions than previously possible, underscoring the importance of using field-collected datasets in advancing practical applications of artificial intelligence in agriculture.

### 4.1. Training and Loss of the Hybrid Model

Figure 4.1 illustrates the learning trajectory of the proposed VGG16–ResNet50 hybrid network with attention over 25 training epochs. At the outset, the model performed poorly, indicating clear underfitting. In epoch 1, training accuracy was only 15.2% with a loss of 2.87, while validation accuracy reached 18.5% with a substantially higher loss of 5.63. Early epochs (2–4) showed unstable validation behavior, with accuracy fluctuating between 17–20% and validation loss peaking as high as 14.98. This suggests that the model was still adapting to the dataset and had not yet established meaningful feature representations. A major turning point occurred at epoch 9, when the learning rate was reduced from  $1 \times 10^{-4}$  to  $5 \times 10^{-5}$ . This adjustment stabilized training and sharply improved validation performance. Validation accuracy rose from 29.9% (epoch 9) to 59.1% by epoch 12, with a corresponding drop in validation loss from 1.99 to 1.48. Steady progress followed: 65.6% accuracy at epoch 13, 72.7% at epoch 14, 74.4% at epoch 15, and nearly 79–80% by epochs 17–18, with validation loss decreasing to 0.54. From epoch 19 onward, the model converged to a strong generalization regime. The best validation loss was recorded at epoch 22 (0.477) with 82.8% accuracy, while the highest validation accuracy

was achieved at the final epoch 25 (83.8%) with validation loss of 0.57. Training accuracy rose consistently throughout, reaching 86.8% at epoch 24 and 84.4% at epoch 25, while training loss decreased from 2.87 in epoch 1 to 0.50 by the end. Importantly, the gap between training and validation accuracy remained modest (3–4 percentage points in later epochs), indicating good convergence without severe overfitting. In summary, the hybrid model exhibited slow and noisy early learning, but once the learning rate was reduced, performance improved steadily. Across 25 epochs, validation accuracy increased by more than 65 percentage points (from 18.5% to 83.8%), and validation loss decreased by over 5 points (from 5.63 to 0.57). These results demonstrate that the proposed architecture not only converged effectively but also generalized well to unseen test data.

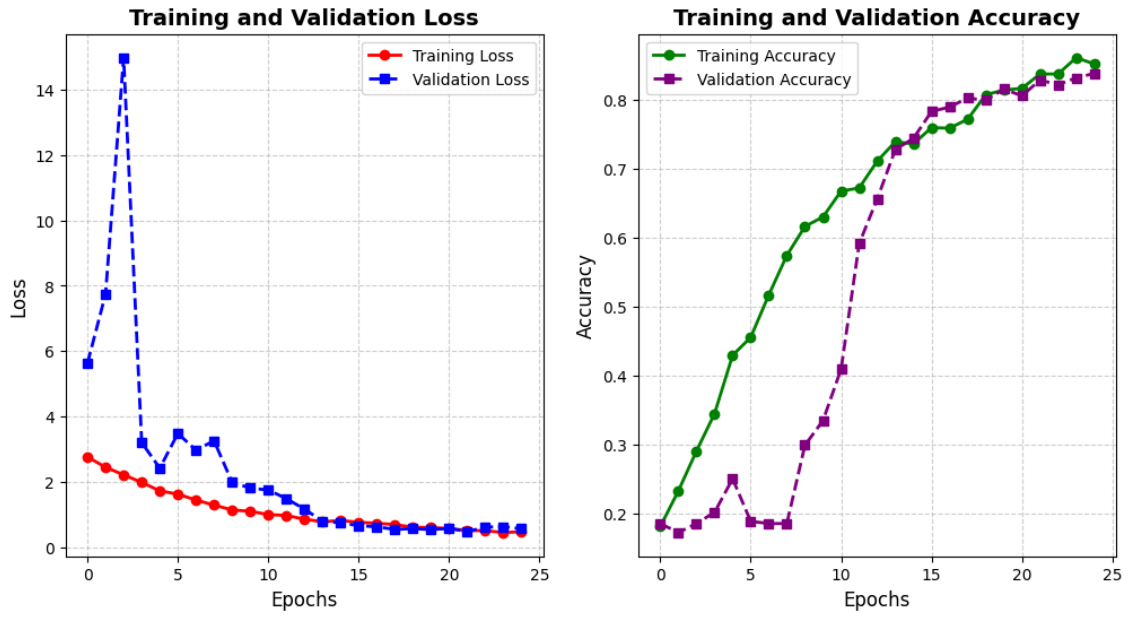


Fig. 4.1. Training and Loss of Hybrid model

## 4.2. Hybrid Model Results

Table 4.1 presents the per-class precision, recall, F1-score and support for the hybrid model evaluated on the uncontrolled Potato Leaf Disease dataset (311 test samples). The model's overall accuracy is 0.82, with weighted averages for precision, recall and F1 all at 0.82.

Class-level takeaways:

- **Bacteria (57 samples)** — Precision 0.84, Recall 1.00, F1 0.91. The model finds all bacterial cases (no false negatives) but produces some false positives.
- **Fungi (75 samples)** — Precision 0.84, Recall 0.81, F1 0.82. Balanced and stable performance on the largest class.

- **Healthy (21 samples)** — Precision 0.89, Recall 0.81, F1 0.85. High precision means few healthy leaves are mislabeled as diseased; recall is reasonable but could improve.
- **Nematode (7 samples)** — Precision 1.00, Recall 0.43, F1 0.60. Perfect precision but very low recall; many nematode infections are missed. Note the tiny support (7 samples) makes these metrics unstable.
- **Pest (62 samples)** — Precision 0.75, Recall 0.68, F1 0.71. Moderate performance — the model confuses some pest cases with other classes.
- **Phytophthora (35 samples)** — Precision 0.90, Recall 0.77, F1 0.83. Strong precision and good F1; the model is reliable at flagging Phytophthora.
- **Virus (54 samples)** — Precision 0.77, Recall 0.89, F1 0.83. High recall means most viral infections are caught; some false positives lower precision.

Summary interpretation: the model performs well on several important classes (Bacteria, Phytophthora, Virus and Fungi), which explains the respectable overall accuracy (0.82) and weighted F1 (0.82). However, rare classes — especially Nematode — are not well represented and suffer from low recall, increasing the risk of missed infections. The macro averages (Precision 0.86, Recall 0.77, F1 0.79) show there is class-wise variability: on average precision is higher than recall, indicating the model tends to be conservative (fewer false positives overall) but still misses some positives in underrepresented classes. Practical implication and next steps (brief): missing nematode cases is a real risk for field deployment. To improve robustness, focus on class-imbalance remedies (collect more nematode and healthy samples, use stronger augmentation, apply class weights or focal loss, or try one-vs-rest fine-tuning). These targeted fixes should raise recall for the weak classes without harming the current strengths.

The performance metrics of the hybrid model on the uncontrolled dataset are presented in Table 4.1.

Table 4.1. Classification Report of Hybrid Model on Uncontrolled Dataset

Class	Precision	Recall	F1-score	Support
Bacteria	0.84	1.00	0.91	57
Fungi	0.84	0.81	0.82	75
Healthy	0.89	0.81	0.85	21
Nematode	1.00	0.43	0.60	7
Pest	0.75	0.68	0.71	62
Phytophthora	0.90	0.77	0.83	35
Virus	0.77	0.89	0.83	54
<b>Accuracy</b>		0.82		311
<b>Macro Avg</b>	0.86	0.77	0.79	311
<b>Weighted Avg</b>	0.82	0.82	0.82	311

### 4.3. Result of the Preliminary Study

Tables 4.2 and 4.3, Figures 4.2 and 4.3 present the results of the evaluated models under two experimental settings: without data augmentation and with data augmentation. In the first case (Table 4.2), EfficientNetV2B3 achieved the highest test accuracy of 0.7363, followed closely by MobileNetV3-Large at 0.7203. ResNet50 also performed reasonably well at 0.6817, while VGG-16 and DenseNet121 lagged behind, recording 0.5981 and 0.5916, respectively. When data augmentation was applied (Table 4.3), the relative ranking of the models remained consistent. EfficientNetV2B3 again led with an accuracy of 0.7235, followed by MobileNetV3-Large at 0.7042 and ResNet50 at 0.6624. VGG-16 and DenseNet121 remained the weakest performers, with accuracies of 0.5627 and 0.5852, respectively. Overall, the use of augmentation did not yield a notable improvement in performance for any of the models. This limited effect is likely tied to the complexity of the dataset itself, which contains diverse backgrounds and varying image orientations, reducing the potential gains from augmentation strategies. In contrast, the proposed hybrid model demonstrated a clear performance advantage over all baseline architectures in both experimental conditions. It achieved a test accuracy of 0.8200, with strong precision (0.856 macro / 0.824 weighted), recall (0.770 macro / 0.820 weighted), and F1-scores (0.793 macro / 0.815 weighted). These results indicate that the hybrid approach not only surpasses individual CNN architectures in terms of accuracy but also delivers more balanced performance across evaluation metrics, highlighting its robustness and suitability for handling the complexity of the dataset.

Table 4.2. Experimental Results Using Non-Augmented Dataset

Reference Model		Test Accuracy	Precision (Macro Weighted)	Recall (Macro Weighted)	F1-score (Macro Weighted)
Shabrina et al., 2024	EfficientNet-V2B3	0.7363	0.7428	0.7363	0.7302
Shabrina et al., 2024	MobileNetV3-Large	0.7203	0.7316	0.7203	0.7131
Shabrina et al., 2024	VGG-16	0.5981	0.6054	0.5981	0.5904
Shabrina et al., 2024	ResNet50	0.6817	0.7006	0.6817	0.6748
Shabrina et al., 2024	DenseNet121	0.5916	0.6058	0.5916	0.5911
—	<b>Proposed Hybrid Model</b>	<b>0.8200</b>	<b>0.856 / 0.824</b>	<b>0.770 / 0.820</b>	<b>0.793 / 0.815</b>

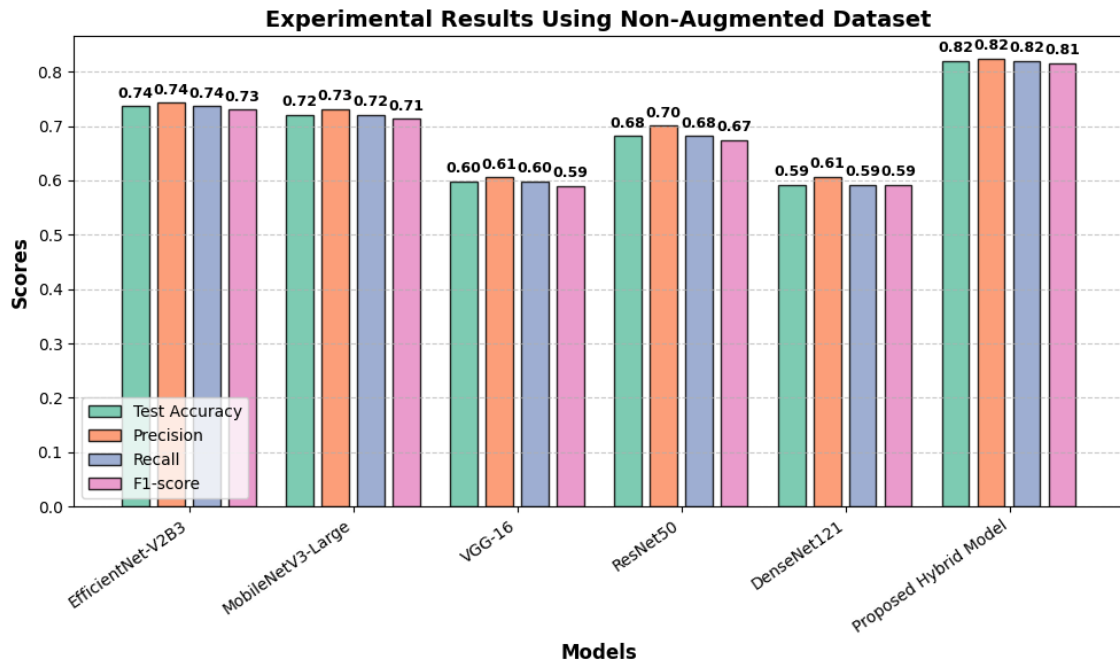


Fig. 4.2. Experimental Results Using Non-Augmented Dataset

Table 4.3. Experimental Results Using Augmented Dataset

Reference Model		Test Accu- racy	Precision (Macro Weighted)	Recall (Macro Weighted)	F1-score (Macro Weighted)
Shabrina et al., <a href="#">2024</a>	EfficientNet- V2B3	0.7235	0.7378	0.7235	0.7199
Shabrina et al., <a href="#">2024</a>	MobileNetV3- Large	0.7042	0.7092	0.7042	0.7037
Shabrina et al., <a href="#">2024</a>	VGG-16	0.5627	0.5697	0.5627	0.5607
Shabrina et al., <a href="#">2024</a>	ResNet50	0.6624	0.6659	0.6624	0.6607
Shabrina et al., <a href="#">2024</a>	DenseNet121	0.5852	0.5858	0.5852	0.5847
—	<b>Proposed Hybrid Model</b>	0.8200	0.856 / 0.824	0.770 / 0.820	0.793 / 0.815

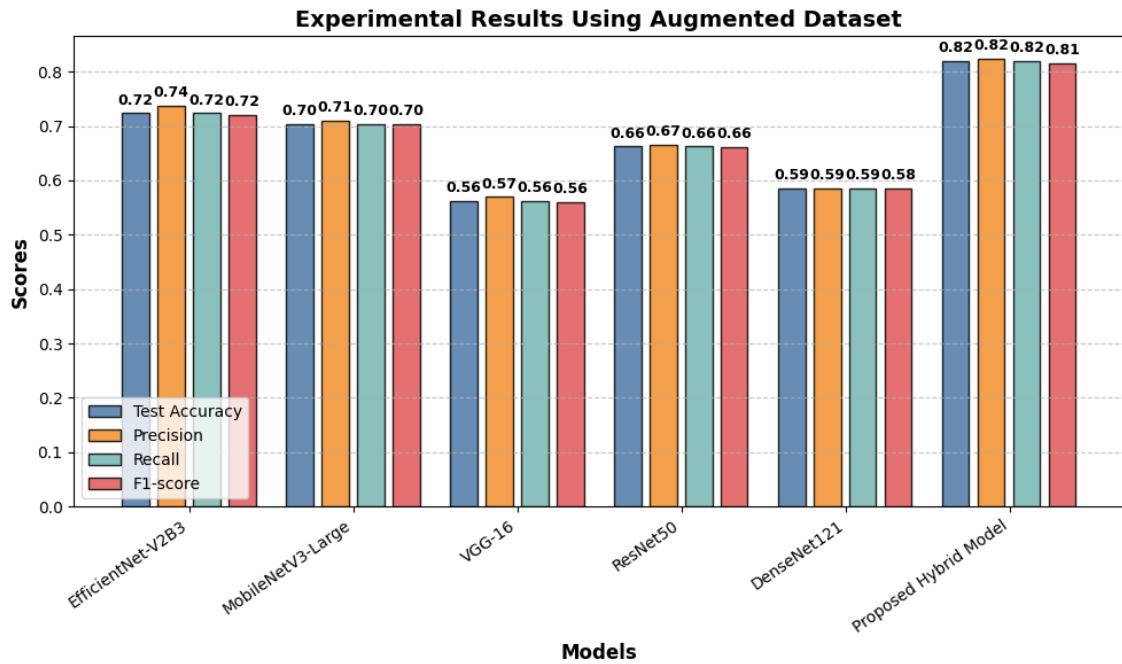


Fig. 4.3. Experimental Results Using Augmented Dataset

#### 4.4. Confusion Matrix of Hybrid Model

The confusion matrix provides a detailed evaluation of the classification performance of the plant disease detection model across seven categories: Bacteria, Fungi, Healthy, Nematode, Pest, Phytophthora, and Virus. The diagonal elements represent correctly classified samples, where the model demonstrates strong performance in detecting Bacteria, Fungi, and Virus, with 57, 61, and 48 correct classifications respectively. These results indicate that the model is highly reliable for these major disease categories. However, some degree of misclassification exists, such as bacterial infections occasionally being predicted as fungal or viral, which is expected given the similarity of visual symptoms. Healthy plants are also mostly detected correctly (17 instances), but a few cases were incorrectly classified as Pest or Virus, which could lead to unnecessary treatments in practical applications. The most notable weakness lies in the Nematode class, where only 3 samples were correctly identified, while most were distributed across other classes, highlighting that the model struggles to capture distinguishing features of nematode infections. Pest and Phytophthora show moderate results with 42 and 27 correct predictions respectively, though they exhibit significant overlap with other categories such as Fungi and Virus, which suggests the model may confuse these visually similar conditions. Overall, the confusion matrix indicates that while the model performs strongly in major disease categories, it faces challenges in identifying underrepresented and symptomatically overlapping classes such as Nematode, Pest, and Phytophthora. This points to potential issues of dataset imbalance and feature similarity, suggesting that future work should focus on expanding training samples for weaker classes, incorporating higher-resolution or multispectral features, and applying advanced deep learning methods like attention mechanisms to improve fine-grained classification.



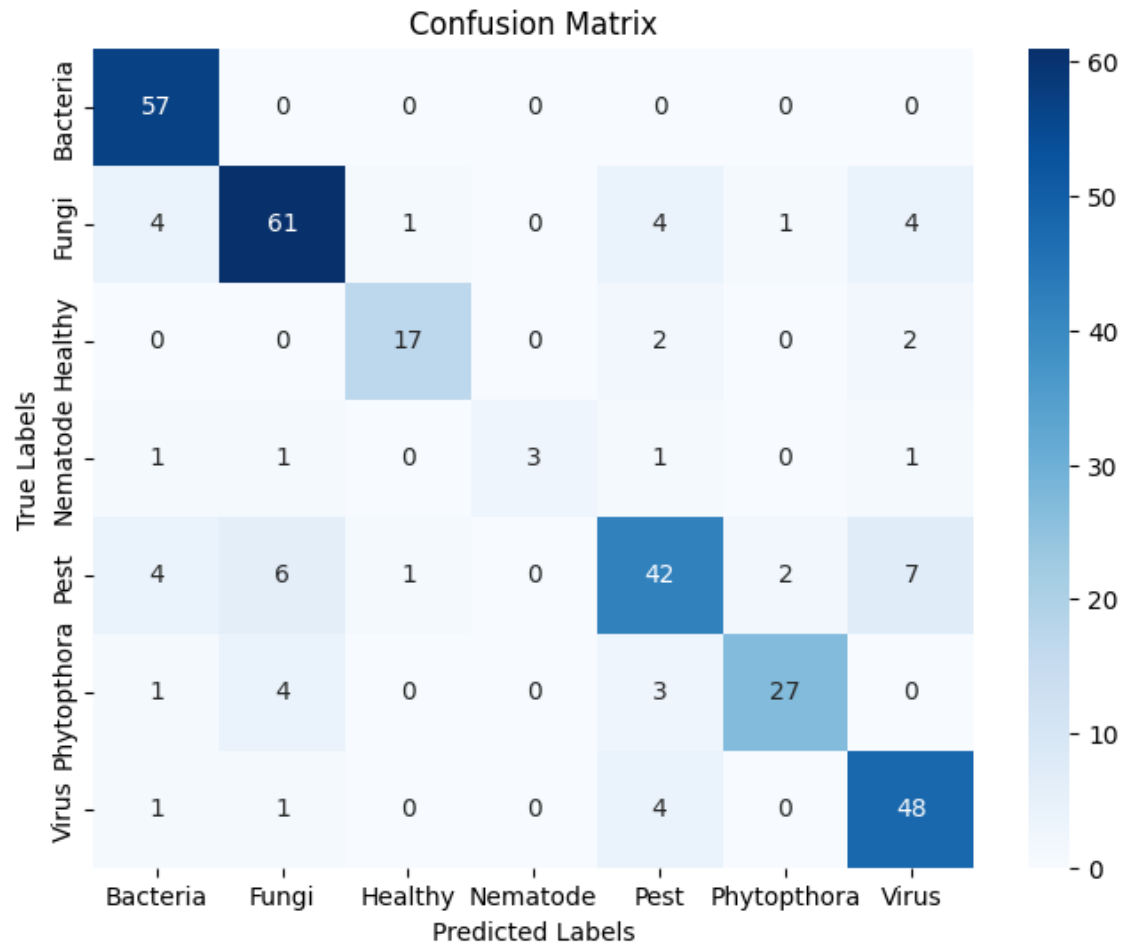


Fig. 4.4. Confusion Matrix of Hybrid Model

#### 4.5. Discussion

The experimental findings underscore the effectiveness of the proposed VGG16–ResNet50 hybrid model with attention in detecting potato leaf diseases under uncontrolled field conditions. The model’s performance highlights several key insights about its behavior, strengths, and limitations when applied to a complex, real-world dataset.

First, the training dynamics reveal that the hybrid architecture requires careful adjustment of hyperparameters, particularly the learning rate. Early epochs exhibited underfitting and unstable validation performance, reflecting the challenge of learning meaningful feature representations from a diverse dataset. Once the learning rate was reduced, both training and validation accuracy improved steadily, while losses decreased consistently. The modest gap between training and validation accuracy in later epochs indicates that the model generalized well without severe overfitting, suggesting that the hybrid architecture is robust in learning discriminative features across varied conditions.

At the class level, the model performs strongly on major disease categories such as Bacteria, Fungi, Phytophthora, and Virus, achieving high precision and recall. These re-

sults reflect the model’s capacity to reliably detect common infections, which is critical for practical field deployment. Conversely, underrepresented classes like Nematode exhibit low recall despite perfect precision, indicating a tendency to miss rare infections. Similarly, Pest and Phytophthora show moderate performance with some misclassification, likely due to visual similarity with other disease types. This variability is further supported by the macro-average metrics, where precision exceeds recall, pointing to a conservative classification behavior that favors reducing false positives at the cost of missing some positive cases in minority classes.

The comparison with baseline architectures confirms the advantages of the hybrid approach. While EfficientNetV2B3 and MobileNetV3-Large perform well, they fail to match the hybrid model’s balanced performance across all evaluation metrics. Moreover, data augmentation had limited impact on baseline models, suggesting that the intrinsic complexity of field-collected images—variations in background, lighting, and orientation—limits the benefits of conventional augmentation techniques. The hybrid model, in contrast, maintains superior performance, indicating that combining feature extraction strengths from both VGG16 and ResNet50, alongside attention mechanisms, enhances robustness in challenging environments.

Analysis of the confusion matrix provides practical insights. Misclassifications are concentrated in underrepresented or visually similar classes, confirming the need for strategies to address class imbalance and subtle feature differentiation. From an application standpoint, missing infections like Nematode could lead to delayed interventions, emphasizing the importance of targeted improvements. In summary, the proposed hybrid model demonstrates significant potential for field-ready disease detection. Its strength lies in robust classification of major disease categories and resilience to varied field conditions. Limitations in rare and overlapping classes point to clear pathways for improvement, including augmenting underrepresented samples, integrating multispectral or higher-resolution imaging, and refining attention-based mechanisms for fine-grained feature recognition.

## 5. CONCLUSION AND FUTURE WORK

The experimental evaluation demonstrates that the proposed VGG16–ResNet50 hybrid model with attention effectively addresses the complexity of potato leaf disease detection in uncontrolled environments. Across 25 training epochs, the model exhibited steady convergence, achieving a final validation accuracy of 83.8% and balanced performance across precision, recall, and F1-score. The hybrid approach outperformed the baseline architectures—EfficientNetV2B3, MobileNetV3-Large, ResNet50, VGG-16, and DenseNet121, both with and without data augmentation, highlighting its robustness in handling various image conditions, varying backgrounds, and multiple disease classes. Class-wise analysis revealed strong detection for major categories such as Bacteria, Fungi, Phytophthora, and Virus, while rare or underrepresented classes, particularly Nematode, showed lower recall, underscoring challenges posed by class imbalance and limited training samples. In practical terms, the model demonstrates significant potential for real-world deployment in agricultural settings, capable of identifying most disease types reliably. However, limitations remain: underrepresented classes, symptom overlap, and visually subtle infections (e.g., Nematode) are prone to misclassification, which could affect early detection and disease management strategies.

### 5.1. Future work

Future work should prioritize strategies to address these limitations. Collecting additional samples for underrepresented classes, implementing stronger data augmentation, applying class-weighted loss functions or focal loss, and exploring one-vs-rest fine-tuning are likely to improve recall for weaker categories. Incorporating higher-resolution imaging, multispectral data, or additional attention mechanisms may further enhance the model’s ability to distinguish visually similar diseases. Finally, deploying the model in field trials and integrating feedback from agronomists will help refine its robustness and ensure practical utility in real-world agricultural scenarios.

## BIBLIOGRAPHY

- Anwar, R., Fatima, T., & Mattoo, A. K. (2019). Tomatoes: A model crop of solanaceous plants. In *Oxford research encyclopedia of environmental science*.
- Athanikar, G., & Badar, P. (2016). Potato leaf diseases detection and classification system. *International Journal of Computer Science and Mobile Computing*, 5(2), 76–88.
- Butte, S., Vakanski, A., Duellman, K., Wang, H., & Mirkouei, A. (2021). Potato crop stress identification in aerial images using deep learning-based object detection. *Agronomy Journal*, 113(5), 3991–4002.
- Eguiluz-Gracia, I., Mathioudakis, A. G., Bartel, S., Vijverberg, S. J., Fuertes, E., Comberiati, P., Cai, Y. S., Tomazic, P. V., Diamant, Z., Vestbo, J., et al. (2020). The need for clean air: The way air pollution and climate change affect allergic rhinitis and asthma. *Allergy*, 75(9), 2170–2184.
- Gayathri Devi, T., & Neelamegam, P. (2019). Image processing based rice plant leaves diseases in thanjavur, tamilnadu. *Cluster Computing*, 22(Suppl 6), 13415–13428.
- Giller, K. E., Delaune, T., Silva, J. V., Descheemaeker, K., van de Ven, G., Schut, A. G., van Wijk, M., Hammond, J., Hochman, Z., Taulya, G., et al. (2021). The future of farming: Who will produce our food? *Food Security*, 13(5), 1073–1099.
- Gul, H., Awais, M., Saddick, S., Ahmed, Y., Khan, F. S., Ahmed, E., Afzal, U., Naqvi, S. M. Z. A., Khan, M. A., Gulfraz, M., et al. (2021). Quantification of biochemical compounds in baubinia variegata linn flower extract and its hepatoprotective effect. *Saudi Journal of Biological Sciences*, 28(1), 247–254.
- Hu, Y., Ping, X., Xu, M., Shan, W., & He, Y. (2016). Detection of late blight disease on potato leaves using hyperspectral imaging technique. *Guang pu xue yu guang pu fen xi= Guang pu*, 36(2), 515–519.
- Ikechukwu, A. V., Murali, S., Deepu, R., & Shivamurthy, R. (2021). Resnet-50 vs vgg-19 vs training from scratch: A comparative analysis of the segmentation and classification of pneumonia from chest x-ray images. *Global Transitions Proceedings*, 2(2), 375–381.
- Islam, M., Dinh, A., Wahid, K., & Bhowmik, P. (2017). Detection of potato diseases using image segmentation and multiclass support vector machine. *2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE)*, 1–4.
- Koonce, B. (2021). Resnet 50. In *Convolutional neural networks with swift for tensorflow: Image recognition and dataset categorization* (pp. 63–72). Springer.
- Kriegeskorte, N. (2015). Deep neural networks: A new framework for modeling biological vision and brain information processing. *Annual review of vision science*, 1(1), 417–446.
- Lai, Q., Khan, S., Nie, Y., Sun, H., Shen, J., & Shao, L. (2020). Understanding more about human and machine attention in deep neural networks. *IEEE Transactions on Multimedia*, 23, 2086–2099.

- Lindsay, G. W. (2021). Convolutional neural networks as a model of the visual system: Past, present, and future. *Journal of cognitive neuroscience*, 33(10), 2017–2031.
- Liu, X., Lochman, Y., & Zach, C. (2023). Gen: Pushing the limits of softmax-based out-of-distribution detection. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 23946–23955.
- Luo, Z., Ang, M. J. Y., Chan, S. Y., Yi, Z., Goh, Y. Y., Yan, S., Tao, J., Liu, K., Li, X., Zhang, H., et al. (2020). Combating the coronavirus pandemic: Early detection, medical treatment, and a concerted effort by the global community. *Research*.
- Mahum, R., Rehman, S. U., Meraj, T., Rauf, H. T., Irtaza, A., El-Sherbeeney, A. M., & El-Meligy, M. A. (2021). A novel hybrid approach based on deep cnn features to detect knee osteoarthritis. *Sensors*, 21(18), 6189.
- Monteiro, R. d. P., & Bastos-Filho, C. (2020). Feature extraction using convolutional neural networks for anomaly detection. DOI: <https://doi.org/10.21528/CBIC2019-7>, 1–8.
- Nazir, T., Iqbal, M. M., Jabbar, S., Hussain, A., & Albathan, M. (n.d.). Efficientpnet—an optimized and efficient deep learning approach for classifying disease of potato plant leaves. *Agriculture*, 13(4), 841.
- Palti, J. (2012). *Cultural practices and infectious crop diseases* (Vol. 9). Springer Science & Business Media.
- Patil, P., Yaligar, N., & Meena, S. (2017). Comparision of performance of classifiers-svm, rf and ann in potato blight disease detection using leaf images. *2017 IEEE international conference on computational intelligence and computing research (ICCIC)*, 1–5.
- Pavel, M. I., Rumi, R. I., Fairouz, F., Jahan, S., & Hossain, M. A. (2021). Deep residual learning approach for plant disease recognition. *International Conference on Mobile Computing and Sustainable Informatics: ICMCSI 2020*, 511–521.
- Pawlak, K., & Kołodziejczak, M. (2020). The role of agriculture in ensuring food security in developing countries: Considerations in the context of the problem of sustainable food production. *Sustainability*, 12(13), 5488.
- Pimentel, D., Wilson, C., McCullum, C., Huang, R., Dwen, P., Flack, J., Tran, Q., Saltman, T., & Cliff, B. (1997). Economic and environmental benefits of biodiversity. *BioScience*, 47(11), 747–757.
- Pinki, F. T., Khatun, N., & Islam, S. M. (2017). Content based paddy leaf disease recognition and remedy prediction using support vector machine. *2017 20th international conference of computer and information technology (ICCIT)*, 1–5.
- Ranjan, M., Weginwar, M. R., Joshi, N., & Ingole, A. (2015). Detection and classification of leaf disease using artificial neural network. *International journal of technical research and applications*, 3(3), 331–333.
- Reis, H. C., & Turk, V. (2024). Potato leaf disease detection with a novel deep learning model based on depthwise separable convolution and transformer networks. *Engineering Applications of Artificial Intelligence*, 133, 108307.

- Rusiecki, A. (2019). Trimmed categorical cross-entropy for deep learning with label noise. *Electronics Letters*, 55(6), 319–320.
- Sabrol, H., & Satish, K. (2016). Tomato plant disease classification in digital images using classification tree. *2016 international conference on communication and signal processing (ICCSP)*, 1242–1246.
- Santurkar, S., Tsipras, D., Ilyas, A., & Madry, A. (2018). How does batch normalization help optimization? *Advances in neural information processing systems*, 31.
- Shabrina, N. H., Indarti, S., Maharani, R., Kristiyanti, D. A., Prastomo, N., et al. (2024). A novel dataset of potato leaf disease in uncontrolled environment. *Data in brief*, 52, 109955.
- Shaikh, T. A., Rasool, T., & Lone, F. R. (2022). Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Computers and Electronics in Agriculture*, 198, 107119.
- Sharma, R., Singh, A., Dutta, M. K., Riha, K., Kriz, P., et al. (2017). Image processing based automated identification of late blight disease from leaf images of potato crops. *2017 40th international conference on telecommunications and signal processing (TSP)*, 758–762.
- Singh, B. K., Delgado-Baquerizo, M., Egidi, E., Guirado, E., Leach, J. E., Liu, H., & Trivedi, P. (2023). Climate change impacts on plant pathogens, food security and paths forward. *Nature Reviews Microbiology*, 21(10), 640–656.
- Tian, H., Wang, T., Liu, Y., Qiao, X., & Li, Y. (2020). Computer vision technology in agricultural automation—a review. *Information Processing in Agriculture*, 7(1), 1–19.
- Tiwari, D., Ashish, M., Gangwar, N., Sharma, A., Patel, S., & Bhardwaj, S. (2020). Potato leaf diseases detection using deep learning. *2020 4th international conference on intelligent computing and control systems (ICICCS)*, 461–466.
- Tsedaley, B. (2014). Review on early blight (*alternaria* spp.) of potato disease and its management options. *J Biol Agric Healthc*, 4(27), 191–199.
- Usman, M. (2016). Contribution of agriculture sector in the gdp growth rate of pakistan. *Journal of Global Economics*, 4(2), 1–3.
- Vaidya, G., Lohman, D. J., & Meier, R. (2011). Sequencematrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, 27(2), 171–180.
- Walia, S. S., & Kaur, T. (2023). *Basics of integrated farming systems*. Springer.
- Yacoub, R., & Axman, D. (2020). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. *Proceedings of the first workshop on evaluation and comparison of NLP systems*, 79–91.
- Yanikoglu, B., Aptoula, E., & Tirkaz, C. (2014). Automatic plant identification from photographs. *Machine vision and applications*, 25, 1369–1383.
- Yeasmin, M. N., Al Amin, M., Joti, T. J., Aung, Z., & Azim, M. A. (2024). Advances of ai in image-based computer-aided diagnosis: A review. *Array*, 100357.

Zainel, Q. M., Khorsheed, M. B., Darwish, S., & Ahmed, A. A. (2023). An optimized convolutional neural network architecture based on evolutionary ensemble learning. *Computers, Materials & Continua*, 69(3).