Cucumber Leaf Disease Detection and Classification Using Artificial Intelligence

By

Somaya Sultana
ID: CSE2103024016

Ayesha Siddika Asha
ID: CSE2201025045

Sonya Chowdhury
ID: CSE2202026045

Tanvirujjaman
ID: CSE2202026130

Supervised By:
Salma Tabashum

Submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science & Engineering.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SONARGAON UNIVERSITY (SU)

May 2026

APPROVAL

The thesis titled "**Cucumber Leaf Disease Detection and Classification Using Artificial Intelligence**" submitted by Somaya Sultana (CSE2103024016), Ayesha Siddika Asha (CSE2201025045), Sonya Chowdhury (CSE2202026045) and Tanvirujjaman (CSE2202026130) to the Department of Computer Science and Engineering, Sonargaon University (SU), has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering and approved as to its style and contents.

Board of Examiners

Salma Tabashum **Supervisor**
Lecturer & Coordinator
Department of Computer Science and Engineering
Sonargaon University (SU)

(Examiner Name and Signature) **Examiner 1**
Department of Computer Science and Engineering
Sonargaon University (SU)

(Examiner Name and Signature) **Examiner 2**
Department of Computer Science and Engineering
Sonargaon University (SU)

--------------------------------------- **Examiner 3**

(Examiner Name and Signature)
Department of Computer Science and Engineering
Sonargaon University (SU)

DECLARATION

We, hereby, declare that the work presented in this report is the outcome of the investigation performed by us under the supervision of **Salma Tabashum**, **Lecturer** and **Coordinator,** Department of Computer Science and Engineering, Sonargaon University, Dhaka, Bangladesh. We reaffirm that no part of this thesis has been or is being submitted elsewhere for the award of any degree or diploma.

Countersigned	Signature
------------------------------	------------------------
(Salma Tabashum)	Somaya Sultana
Supervisor	ID: CSE2103024016

	Ayesha Siddika Asha
	ID: CSE2201025045

	Sonya Chowdhury
	ID: CSE2202026045

	Tanvirujjaman
	ID: CSE2202026130

ABSTRACT

Cucumber (*Cucumis sativus* L.) is a vital vegetable crop in Bangladesh, highly valued for its nutritional benefits and economic significance. However, cucumber cultivation is severely threatened by various leaf pathogens such as powdery mildew, downy mildew, anthracnose, and bacterial wilt; which drastically reduce crop yield and quality. Early and precise disease identification is critical for implementing targeted interventions and minimizing agricultural losses.

This thesis presents an advanced, Artificial Intelligence (AI)-based diagnostic system for the automated detection and classification of cucumber leaf diseases. The proposed framework evaluates a custom Convolutional Neural Network (CNN), standard transfer learning architectures (MobileNetV2, ResNet50, and EfficientNetB0), and a hybrid deep learning model combining EfficientNet with a Support Vector Machine (EfficientNet-SVM). To train these models, a comprehensive dataset of labeled leaf images was compiled from both online repositories and real-field environments in Bangladesh, and subsequently optimized using resizing, normalization, and data augmentation techniques.

The models were rigorously evaluated using standard performance metrics, including accuracy, precision, and recall. Experimental results demonstrate that the proposed hybrid EfficientNet-SVM model achieved the highest overall performance, yielding an exceptional 99.09% accuracy, 99.00% precision, and 99.00% recall. Among the end-to-end deep learning architectures, MobileNetV2 and the Custom CNN performed robustly (achieving 96.17% and 95.96% accuracy, respectively), while traditional ensemble methods like Random Forest (19.33%) and deeper architectures like ResNet50 (60.57%) struggled to generalize.

The superior performance of the hybrid model, alongside the lightweight footprint of MobileNetV2, highlights the immediate viability of deploying these systems via mobile applications. This research offers a transformative tool for Bangladeshi farmers to rapidly diagnose crop diseases, reduce chemical pesticide overuse through precise targeting, and foster sustainable, technology-driven agricultural practices.

ACKNOWLEDGMENT

At the very beginning, we would like to express my deepest gratitude to the Almighty Allah for giving us the ability and the strength to finish the task successfully within the scheduled time.

We are auspicious that we had the kind association as well as supervision of **Salma Tabashum**, Lecturer & Coordinator, Department of Computer Science and Engineering, Sonargaon University, whose heartfelt and valuable support with the best concern and direction acted as a necessary recourse to carry out our project.

We would like to convey our special gratitude to Prof. **Bulbul Ahamed**, Head, Department of Computer Science and Engineering, for his kind concern and precious suggestions.

We are also thankful to all our teachers during our whole education for exposing us to the beauty of learning.

Finally, our deepest gratitude and love to my parents for their support, encouragement, and endless love.

LIST OF ABBREVIATIONS

A-Hash	Average Hash
AI	Artificial Intelligence
CNNs	Convolutional Neural Networks
D-Hash	Difference Hash
DCN	Deformable Convolutional Networks
DL	Deep Learning
FLOP	Floating Point Operation
FPS	Frames Per Second
GLCM	Grey Level Co-occurrence Matrix
GPDCNN	Global Pooling Deep Convolutional Neural Network
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
IoT	Internet of Things
KNN	k-Nearest Neighbor
LSTM	long short-term memory
MBConv	Mobile Inverted Bottleneck Convolution
MD5 Hash	Message Digest Algorithm 5
ML	Machine Learning
P-Hash	Perceptual Hash
ReLU	Rectified Linear Unit
SE	squeeze-and-excitation
SNN	Spiking Neural Network
SVM	Support Vector Machine
U-Net	U-shaped Convolutional Neural Network
VGG16	Visual Geometry Group 16
ViT	Vision Transformer
YOLO	You Only Look Once

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION ABOUT THIS THESIS

1.1 Introduction

Bangladesh's economy depends heavily on agriculture. It helps rural communities and supplies food for a large number of people. Vegetable farming is a significant component of this sector. Cucumis sativus, the scientific name for cucumber, is one of the most widely grown vegetables worldwide. It grows rapidly, is nutritious, and has a high market value. It does, however, face a number of difficulties. Cucumbers are susceptible to a number of pests and illnesses, such as leaf spot, powdery mildew, and downy mildew. These issues may reduce yields and have an impact on the produce's appearance.

Farmers often stick to traditional methods for growing their crops. They check plants by hand for pests, which can take a lot of time. Additionally, this method is not always reliable, even for skilled farmers. If they miss something or take too long to notice a problem, they may use too much pesticide. This can hurt the environment and lead to higher costs.

Researchers have developed new technologies that use artificial intelligence, including deep learning and image processing. These advancements involve Convolutional Neural Networks (CNNs), which can spot specific diseases on plant leaves. CNNs can pull complex details from images and accurately identify plant issues. A CNN can identify any diseases affecting a cucumber plant when a farmer takes a picture of the leaf.

This research could help Bangladeshi farmers improve their crops. It can lead to healthier plants, better access to markets, and support for sustainable farming practices.

1.2 Why is it important?

The study "Cucumber Leaf Disease Detection and Classification Using Artificial Intelligence" draws attention to the emphasis of several factors containing academic besides agricultural and even economic viewpoints.

Firstly, the fact that cucumbers are the most important vegetable in Bangladesh leads to the country losing a lot in terms of crop yield and quality if leaf diseases are not controlled. Thus, the early detection of these diseases is extremely important to save the crops from the worst scenario. This work presents a method that is both automated and exact for the early detection of cucumber leaf diseases, hence giving the farmers the chance to act and implement the preventive measures in time.

Owing to the above reasons, traditional methods of disease detection have been relying on the manual observations made by either farmers or agricultural experts, which can be extremely costly, labor-intensive, and prone to human errors. An AI-powered imaging classification of diseases comes in with quicker and more accurate diagnostics, consequently lessening the dependence on expert evaluation. The deep learning methods such as Convolutional Neural Networks (CNNs) that are now levels deeper have been employed by the system to swipe under the microscope the whole leaf images with its untrained eyes and come up with the most important features of the disease without the need for any manual feature extraction. This not only speeds up the classification but also boosts accuracy as compared to traditional methods.

In conclusion, the union of AI and modern farming practices, as illustrated by this research, is a significant step towards smart agriculture in Bangladesh. The system that has been proposed is such that it can easily be transitioned from computer to mobile or web platforms, thus making it ultimately available for the farmers, which in turn, during the process of adoption, will lead to sustainable agricultural progress.

1.3 Objectives of Thesis Work

This dissertation primarily aims to create a system that will use artificial intelligence to recognize and categorize cucumber leaf disease based on images. Its specific objectives include the following:

1. Early and accurate identification of cucumber leaf diseases will enhance crop quality.
2. Timely diagnosis will help increase crop quantity (yield) by avoiding production losses caused by diseases.
3. Farmers will be provided with an efficient and effective solution tool by making agricultural expert. .
4. An accessible system designed for farmers with minimal technical skills will be created.
5. Support for early problem-solving will enable farmers to implement preventive actions before diseases become severe.

1.4 Literature Review

Research scientists focus on plant disease detection because it affects their ability to produce agricultural crops. The existing methods depend on human inspectors who take extended periods to complete their work, which results in mistakes. The introduction of artificial intelligence systems together with deep learning technology, particularly through the development of Convolutional Neural Networks, has brought about major advancements that enhance both accuracy and efficiency in plant disease detection systems.

A. Existing Plant Disease Detection Systems

The initial systems developed for plant disease detection relied on:

- Expert manual observation methods
- Basic image processing techniques which included
- Color segmentation
- Edge detection
- Texture analysis

The system required domain expertise to operate because it had three main drawbacks which included:

- The system had low accuracy
- The system failed to function properly under different lighting conditions
- The system could not handle operations beyond its current capacity
- The system required domain expertise to operate

Machine learning and deep learning methods became the new standard because of these system limitations.

B. Use of CNN in Agriculture

Convolutional Neural Networks (CNNs) are widely used in agricultural applications because they can automatically extract image features through their built-in capabilities.

Why CNN?

- The system eliminates the requirement of human-operated feature discovery work.
- The system acquires knowledge about sophisticated patterns which include color and shape and texture attributes.
- The system demonstrates exceptional accuracy when performing image classification tasks.
- The system uses agricultural applications to solve the following problems:
- The system uses leaf disease detection to identify plant diseases which affect their leaves thus showing their presence.
- The system performs plant analysis through crop identification procedures.
- The system assesses fruit quality through evaluation procedures.
- The system uses weed detection to identify undesired plants that compete with crops.

The use of CNN models shows better performance results compared to conventional techniques which have less accurate and less dependable results.

C. Previous Works on Leaf Disease Classification

Several researchers have used deep learning models to detect plant diseases through their research work.
Research studies which employed basic CNN models showed good accuracy results which ranged from (85 to 92 %) accuracy. The advanced architectural designs of ResNet MobileNet and EfficientNet deliver better results because their networks use extra depth and their systems use improved parameter settings.

Key Findings: Larger datasets improve model performance Data augmentation improves generalization capabilities Transfer learning decreases the duration required for training.
Most studies confirm that deep learning models succeed in detecting diseases at real-time performance.

D. Accuracy Comparison of Models

Different deep learning models have been compared for leaf disease classification:

Model	Accuracy	Features
Traditional ML	70–85%	Manual features
CNN	85–95%	Automatic feature extraction
ResNet	90–97%	Deep architecture, high accuracy
MobileNet	88–96%	Lightweight, fast
EfficientNet	92–98%	Optimized scaling

Table 1.4.4 Accuracy Comparison of Models

1.5 Real Life Applications

- **Early Disease Detection on Farms:**

The AI-based solution that has been proposed is capable of scanning cucumber fields directly for precocious leaf disease detection. Precocious identification enables farmers to mitigate damages turned to crops and, thus, control the disease spread by taking timely measures.

- **Facing Mobile Applications:**

Provide offline smartphone access to instant disease scanning and treatment recommendations and local agronomist referral services.

- **Extension Worker Dashboards:**

Display geotagged disease heatmaps which enable regional disease monitoring and resource allocation.

- **Drone & IoT Integration:**

Automated canopy scanning to control variable-rate sprayers or ventilation systems on large farms and greenhouses.

- **Agri-Input & Supply Chain Optimization:**

Data-driven pesticide distribution and harvest timing prediction and crop insurance risk assessment combine to create a system that optimizes.

- **Educational & Capacity-Building Platforms:**

Provide interactive diagnostic tools which support farmer training programs and agricultural curriculum development.

1.6 How does it work?

The AI-based cucumber leaf disease detection system operates through a structured, multi-stage pipeline:

Digital Image Capture: A digital image of a cucumber leaf is captured using a smartphone, field camera, or greenhouse monitoring system. The image shows natural backgrounds with different lighting conditions and parts of the image blocked from view.

Preprocessing: The image is resized to a standard dimension (e.g., 224×224 pixels), color-normalized, and optionally enhanced using techniques like CLAHE or background segmentation to isolate the leaf region.

Feature Extraction & Classification: The preprocessed image is sent to a deep learning model that has been trained on ResNet50 and EfficientNet-B0 and MobileNetV2. The CNN backbone automatically extracts hierarchical visual features through multiple convolutional and pooling layers which include edges and textures and lesion patterns and color distributions.

Prediction Generation: The final fully connected layer outputs a probability distribution across predefined disease classes. The class with the highest confidence score is selected as the predicted diagnosis.

Output & Decision Support: The system returns the disease label, confidence percentage, and optionally displays a heatmap (e.g., Grad-CAM) highlighting the regions influencing the prediction. Integrated advisory modules may suggest targeted treatments or cultural practices.

Deployment Architecture: The model needs to be exported to lightweight formats for field use which include TensorFlow Lite and ONNX and CoreML to make it compatible with mobile devices and edge processors. Cloud-based alternatives enable centralized model updates and continuous learning from user-submitted images.

1.7 Contribution of Thesis Work

The study creates multiple significant advancements for artificial intelligence agricultural research which focuses on cucumber leaf disease identification and classification.

1.7.1 Dataset Development

Researchers built a collection of cucumber leaf images from different sources which included both public databases and actual field photographs. The preprocessing pipeline uses hashing methods to detect duplicate images, which it removes from the dataset to maintain data integrity. The reliable dataset construction which results from this process enhances model training while enabling better generalization.

1.7.2 Enhanced Learning Strategy (Data Augmentation and Transfer Learning)

Advanced learning strategies help to enhance model performance through their implementation. The research team uses data augmentation techniques which include rotation, flipping, zooming, and contrast adjustment to create a more diverse dataset. The team uses transfer learning models MobileNetV2, ResNet50, and EfficientNet B0 to achieve better accuracy while needing less training time.

1.7.3 Lightweight Architecture Evaluation

The research evaluates lightweight deep learning models through its assessment of MobileNetV2 which operates effectively in mobile environments and low-resource settings. The developed system achieves its real-world deployment capability because it meets operational requirements needed for rural farmers.

1.7.4 Comprehensive Performance Analysis

The performance evaluation process requires multiple metrics to assess performance through accuracy, precision, recall, and confusion matrix evaluation. The comparison of different models shows their strengths and limitations, which results in finding the most effective solution.

1.7.5 Farmer-Centric Artificial Intelligence Solution

The proposed system is designed with a focus on practical usability for farmers. The system provides an easy and quick solution that needs no special training to detect cucumber leaf diseases through image analysis. The system helps Bangladesh achieve smart agriculture because it decreases crop loss and increases agricultural productivity.

CHAPTER 2
INTRODUCTION TO CUCUMBER & ITS LEAF DISEASE

2.1 History of Cucumber

The National Garden Bureau has chosen cucumber as its vegetable to promote this year. Thus, 2014 is "**The Year of the Cucumber**". Cool and refreshing on a hot summer day, cucumber is one of America's top five garden vegetables. For years, cucumber was considered to be primarily "**diet food**" because of its low caloric content. Recent research, however, has shown this water-laden vegetable contains significant amounts of phyto-nutrients which have a wide array of human health benefits. While March is a bit early to plant cucumbers in our gardens, it is a fitting month to take a closer look at this interesting vegetable.

Cucumber (Cucumis sativus) belongs to the Cucurbitaceae family. Other important members of this family include watermelon, muskmelon, pumpkin and squash. Native to India, cucumber is another one of our most ancient vegetables. Cave excavations have revealed that cucumber has been grown as a food source for over 3000 years. Early cucumbers were probably very bitter because of compounds they contained called cucurbitacins. These natural defense compounds act to repel insects and other pests. Bitterness still is a problem with some cucumbers today, although great progress has been made by plant breeders to eliminate the bitter compounds.

Cucumbers were cultivated and eaten in ancient Egypt as evidenced by the Bible. Numbers 11:5 reads: "We remember the fish, which we did eat in Egypt freely; the cucumbers, and the melons, and the leeks, and the onions, and the garlic". Evidently the Egyptians made weak liquor from cucumber by cutting a hole in the ripe fruit, stirring the inside with a stick to liquefy it, plugging the hole and then burying it in the ground for several days. The resultant concoction was unearthed and consumed. The previous probably should be followed with the warning: "Don't try this at home".

The Greeks also cultivated cucumbers as (later) did the Romans. The Emperor Tiberius was said to have demanded that they be on his table every day. To meet that demand, his gardeners fashioned some of the first plant forcing structures, constructing portable containers which they moved from place to place. Later, the Romans advanced plant forcing structure technology (and cucumber production) by fashioning frames covered with translucent panes of silicates—not unlike our modern cold frames.

Later, Charlemagne was said to have grown cucumbers in his gardens in Italy in the 8th and 9th century. Cucumber later spread to Western Europe. It was during the reign of King Henry VIII that cucumber made its way to England. His first wife (Catherine of Aragon) was said to have demanded them for her Spanish salads.

The Age of Discovery proved to be very important to the spread of cucumber. Columbus is credited for taking cucumber to the New World, along with many other vegetables. He introduced it to Haiti in 1494. From there it spread. In 1539, De Soto judged the cucumbers he found growing in Florida to be better than those of native land of Spain. By 1806 eight varieties of cucumbers would be found growing in America's colonial gardens.

As mentioned previously, scientists recently have found cucumbers to contain a number of beneficial phyto-nutrients. The lack of that knowledge did not keep cucumber from being used medicinally earlier in our history. The cucumber's water retention ability earned it the reputation for never losing its cool. As a result, 17th Century physicians prescribed placing fever patients on a bed of cucumbers so they would become "cool, as a cucumber." Additionally, it was thought that if cucumbers were eaten three times daily, red noses would be healed and pimples on the face would disappear.

Trappers and fur traders probably were responsible for introducing Native Americans to cucumbers as they made their way across the country. Reportedly, this expansion was interrupted somewhat in the 18th century when medical journals warned of the dangers of eating cucumbers and other vegetables not adequately cooked. The setback was short-lived and by the 19th century cucumber regained its popularity in the U.S.

Cucumber's popularity was bolstered considerably when, in 1876, Henry J. Heinz added pickles to the list of products made by the food processing company that bears his name. Others followed, and by the end of the 19th century pickles were a tasty addition to the monotonous diet of meat and potatoes consumed by most Americans. Today, pickles can be found in 70 percent of all households, and Americans are said to consume nine pounds of them per capita annually.

Cucumbers are placed into one of two major categories based on their use: slicing or pickling. Slicing cucumbers are eaten fresh from the garden. They range in size from about four to 12 inches and have skin that is relatively smooth. If "spines" are present, they normally are white. Pickling cucumbers are much smaller, ranging in size from one to five inches; the latter reserved for making dill pickles. Unlike the slicers, pickling cucumbers usually have bumps on their skin along with "spines" that are black. Gerkins are a type of immature pickling cucumber. They are noted for their small size and warty skin.

Whatever the type, cucumbers contain cucurbitacins which can impart a bitter taste. Additionally, the taste is associated with a digestive condition and social indiscretion known as a "burp". In the middle of the 20th century, "burpless" cucumbers were introduced that eliminated this problem. Today there are several burpless varieties available. Their fruit usually is long, somewhat narrow and thin-skinned. Non-burpless types can be made a bit more socially acceptable by cutting off about one inch of the stem-end and peeling the skin off the fruit.

Cucumbers are relatively easy to grow. They prefer a full-sun exposure in a well-drained garden loam. Amending the soil with compost or other forms of organic matter will improve soil structure and increase yields. Check garden drainage before planting, since wet soil will promote disease infestation and reduce productivity.

Heavy feeders, cucumbers require fertile soil, nitrogen fertilizer, and/or additions of high-nitrogen organic matter sources. A side dressing of 5-10-10 fertilizer at the time of planting and once a month thereafter should be sufficient. Adequate amounts of water also must be provided since the fruits are largely water. Mulching helps to conserve water as well as control weeds. Cucumber plants are vines and need room to grow. Standard types may spread four to six feet and should be spaced four to five feet apart. The restricted vines of dwarf and bush types may require as little as two square feet per plant.

For earlier maturity, cucumbers can be started indoors in individual pots ten to 14 days prior to setting them out. If directly seeded, sow seeds when the soil has warmed up to 70°F. Place one seed every six inches, covering it with soil to a depth of one inch.
Depending on conditions, seedlings should emerge in about a week. When the plants are two inches high, thin plants to one plant per lineal foot. An alternative method is to plant in "hills" four to five feet apart. Sow four to five seeds per hill and thin to three plants per hill after seedlings emerge.

Cucumbers are plagued by a number of troublesome insect pests. They include cucumber beetles (stripped and spotted), aphids and squash vine borers. Control of beetles is important for the prevention of bacterial wilt. Although there are pesticides labeled for the control of the afore-mentioned pest, some gardeners choose to use floating row cover over young transplants or seedlings to exclude the pests.

The two most problematic diseases that afflict cucumber include bacterial wilt and powdery mildew. Bacterial wilt is controlled by eliminating its primary vector—the cucumber beetle. Incidence of powdery mildew can be lessened by providing adequate space between plants, eliminating weeds to increase air circulation, keeping the leaves dry by using drip instead of overhead irrigation, or using fungicides judiciously. Garden clean up in the fall also is helpful in reducing disease inoculums for the following growing season.
Most cucumber varieties reach maturity in 50 to 65 days. The fruit will be firm to the touch and the skin will have a uniform dark green color. Cucumbers enlarge (mature) rapidly and should be checked/harvested daily. If left on the vine too long, the fruit becomes over-mature and less desirable. Slicing types mature when about six to eight inches long; larger slicing varieties should be picked before they are ten inches long. Pickling varieties should be harvested after reaching a length of one to four inches. In either case, refrigerate harvested fruit as soon as possible to preserve flavor and avoid desiccation. [1]

Cucumber is used as:

- Whole fresh
- Sliced fresh and
- Pickled

2.2 Life cycle of cucumber

In general, the life cycle of cucumber in Bangladesh varies from 50 to 70 days, depending on the variety, climate, and cultivation practices. Cucumbers are the most common during the season of winter and summer in Bangladesh.

1.Germinating Stage (3 to 10 days)

Process: Seeds are either cultivated in the field or in seedbeds. The seeds need to absorb moisture and get heated up in order to germinate.

Conditions: The temperature range which is considered as perfect for germination is between 25 and 30 degrees Celsius. The seed within 3 to 10 days will be coming out of the soil as a tiny plant.

2. Seedings Appear (10to 14 days)

Development: The sapling forms its first real leaves (cotyledons).

Root System: The root system starts growing at a fast pace so that it can get water and nutrients from the sandy loam soil which is found in most parts of Bangladesh.

3. Leaves & Vine Grow Larger

Vining: The plant produces vines that may reach a length of 2–3 meters.

Trellising: In Bangladesh, it is the practice to use (trellises) for the vines to climb which in turn leads to the improvement of fruit quality, disease control and easy harvesting.

Nutrient Needs: This phase demands a high amount of nitrogen to be applied for the growth of leaves.

4. Flowering & Pollination stage

Flowers: The plant yields vivid yellow flowers.

Pollination: Generally, the male flowers show up first, then the female ones (which have tiny, underdeveloped fruits at the base).

Pollinators: Bees and various other pollinators play a vital role in the process of pollen being moved from the male to the female flowers.

5. Fruit Formation Stage (Fruiting)

Development: The process of fruit development takes place after pollination has been successful, and the female flowers start to grow fruit.

Water Needs: High humidity and water are necessary for the plant during this stage to avoid the bitterness of the fruit and to ensure a crop of high yield and health.

Duration: The tiny and immature fruits reach their full size within 8-10 days of the process of pollination.

6. Cucumber are Ready to Harvest

The period between day 50 and day 70 of the experiment shows cucumber plants at their actual harvesting point when their fruit reaches its mature stage of firmness and green color at 15-to-20-centimeter length. The plant will produce additional flowers after its mature fruits undergo regular harvesting. A single plant in Bangladesh will produce its harvestable fruit during a period that extends for multiple weeks.

2.3 How to Farm

The climatic conditions in Bangladesh create ideal conditions for cucumber farming which meets the strong market demand for this vegetable. The complete agricultural process needs proper implementation of agricultural methods to achieve better cucumber production results.

1. Land Selection and Preparation

Cucumber farming requires well-drained fertile loamy soil as its ideal growing condition. The land needs multiple plowing sessions to achieve loose soil conditions which will eliminate weed growth. Soil fertility should be improved through the application of organic manure, including compost and cow dung, which will result in higher crop yields and better fruit quality.

2. Seed Selection and Sowing

The selection process demands the use of high-quality certified seeds together with disease-resistant seeds for successful plant growth. Seeds are usually sown during the winter and summer seasons. Proper plant spacing enables improved air flow which decreases disease transmission while enhancing fruit growth.

3. Irrigation Management

Cucumbers need constant irrigation because their growth depends on regular water supply. The use of excessive water needs to be avoided because it creates conditions that lead to root and leaf infections. The correct management of water resources enables plants to maintain their leaf condition which results in increased photosynthetic activity and improved crop production.

4. Fertilizer Application

Balanced use of fertilizers containing nitrogen phosphorus and potassium improves vine growth and leaf development and fruit formation. The application of zinc and boron micronutrients enhances the quality of fruit. Proper fertilization increases both the quantity and market value of cucumbers.

5. Disease and Pest Control

Bangladesh experiences widespread occurrence of leaf diseases which include powdery mildew and downy mildew. The process of early detection requires regular monitoring of leaves. The combination of timely disease management and disease management practices results in reduced crop loss while enhancing the entire production process.

6. Use of Modern Technology

Farmers can detect leaf diseases at their initial stage through the application of artificial intelligence-based disease detection systems which use modern technology. The process of treating crops at their early stage leads to better crop health, higher crop yields, and decreased expenses for farmers.

7. Harvesting and Post-Harvest Care

Cucumbers should be harvested at their optimal maturity stage because this time produces their best flavor and appearance. Proper handling after harvesting reduces damage and maintains quality.

8. Impact on Quality and Quantity

Farmers in Bangladesh can achieve sustainable agricultural production through proper farming techniques and modern AI-based disease detection methods to improve cucumber quality and increase crop yield while decreasing disease-related losses.

2.4 Benefits of Cucumber

2.4.1 Cucumber provides economics and agricultural benefits to Bangladesh.

The cucumber industry offers farmers a profitable investment opportunity because the crop grows and harvests within a short period. The business generates work for agricultural workers and market vendors while creating new wage-earning chances for people in rural areas.

Cucumber serves as an agricultural crop which farmers can grow throughout multiple seasons while using it as part of crop rotation practices. The practice of cucumber cultivation enables farmers to achieve better land productivity while creating an environmentally sustainable method for growing vegetables.

2.4.2 Nutritional and Health Benefits

Component	Concentration (per 100g fresh)	Health Impact
Water	95-96 g	Hydration; low-calorie fullness
Vitamin K	16.4 g (14% DV)	Blood clotting; bone health
Vitamin C	2.8 mg (3% DV)	Antioxidant; immune function
Potassium	147 mg (3% DV)	Electrolyte; heart health
Cucurbitacins	Trace amounts	Anti-inflammatory; possible anti-cancer effects
Dietary fibers	0.5 g	Digestive regulation; blood sugar control

Table 2.4.2 Nutritional and Health Benefits

2.4.3 Environmental Benefits

Water use efficiency, due to the shallow root system, adapted to drip. A shallow root system also implies high WUE with proper water management. Soil health, as a rotation crop, can interrupt certain pest/disease cycles for solanaceous crops. Carbon footprint, short supply chains for local production reduce transport emissions.

2.5 Why cucumbers are important?

Cucumber plays a vital economic role in Bangladesh because it generates revenue for small and marginal farmers while providing job opportunities throughout the process of growing and moving and selling the product. Cucumber enables farmers to achieve fast financial returns because its cultivation period lasts shorter than most other agricultural products. Cucumber farming practices help farmers to create various crops because they use land resources efficiently. Cucumber supports food security and helps people maintain healthy diets according to nutritional standards.

.

2.5.1 Global Food Security Context

Production Scale: >90 million tons annually across 150+ countries (FAO, 2024)

Regional Significance:

Asia: 78% of global production (China: 60 million tons)

Critical protein-sparing vegetable in resource-limited diets

Climate Resilience: Adaptable to diverse Agro-ecologies with appropriate cultivar selection.

2.5.2 Socioeconomic Role

Stakeholder	Importance of Cucumber Cultivation
Smallholder Farmers	Reliable Cash crop; short cycle-aerates income stream; entry barriers cheap.
Women Farmers	Often owns home-garden cucumber plots; adds to household food nutrition and micro-enterprise.
Agri-Processors	Raw material for the pickling, beverage, and cosmetic industries; value-chain jobs.
Urban Consumers	Accessible, high-quality, nutrition vegetable; accessible all year round by protected cultivation.

Table 2.5.2 Socioeconomic Role

2.6 What problem are we facing growing cucumbers? Problems with manual detection

Cucumber production faces permanent restrictions because of foliar diseases which damage its agricultural value. The traditional detection methods create essential obstacles which prevent organizations from achieving their response goals.

2.6.1 Major Cucumber Leaf Diseases

Disease	Pathogen	Primary Symptoms	Yield Impact if Untreated
Powdery Mildew	Podosphaera xanthii	White powdery spots on upper leaf; yellowing beneath; premature defoliation	30–50% loss; reduced fruit size/quality
Downy Mildew	Pseudoperonospora cubensis	Angular yellow lesions on upper surface; fuzzy gray-purple sporulation underneath	40–70% loss; rapid canopy collapse under humid conditions
Angular Leaf Spot	Pseudomonas syringae pv. lachrymans	Water-soaked angular lesions turning tan; may develop "shot-hole" appearance	20–40% loss; fruit scarring reduces marketability
Bacterial Wilt	Erwinia tracheiphila	Wilting of individual leaves/vines; vascular browning; systemic collapse	60–100% loss in susceptible cultivars; no cure post-infection
Target Spot / Corynespora Leaf Spot	Corynespora cassiicola	Circular brown lesions with concentric rings; severe defoliation	25–45% loss; secondary fruit rot

Table 2.6.1 Major Cucumber Leaf Diseases

2.6.2 Limitations of Manual Disease Detection

Challenge	Description	Consequence for Cucumber Farming
Subjectivity & Expertise Gap	Symptom interpretation varies by experience; early-stage lesions are subtle	Misdiagnosis leads to inappropriate treatment (e.g., fungicide for bacterial disease).
Time and Labor Intensity	Scouting 1 hectare requires 2–4 person-hours; weekly monitoring is ideal but rarely feasible	Delayed detection allows exponential pathogen spread; outbreaks become uncontainable.
Spatial Sampling Bias	Farmers inspect easily accessible plants; interior canopy or shaded leaves are overlooked	Hidden infection foci seed secondary outbreaks; field-wide treatment becomes necessary.
Documentation Gaps	Paper-based records lack geotagging, timestamping, or image evidence	Inability to track disease progression, evaluate treatment efficacy, or share data with advisors.
Scalability Constraints	One agronomist can effectively advise 20–30 smallholder farms; demand far exceeds supply	Knowledge asymmetry leaves most farmers reliant on guesswork or input supplier recommendations

Table 2.6.2 Limitations of Manual Disease Detection

CHAPTER 3

INTRODUCTION OF AI & IMPLEMENTATION OF AI IN AGRICULTURE

3.1 What is Artificial Intelligence (AI)?

Artificial intelligence (AI) encompasses the fields of computer and data science focused on building machines with human intelligence to perform tasks like learning, reasoning, problem-solving, perception, and language understanding. Instead of relying on explicit instructions from a programmer, AI systems learn from data that enables them to handle complex problems and simple repetitive tasks, improving how they respond over time.

AI handles many tasks faster with exceptional accuracy and reliability, freeing humans from repetitive and tedious chores. Businesses and organizations save time and money by automating and optimizing routine processes, using the technology to stay connected with customers and gain a competitive edge.

AI is part of everyday life. If you've used a self-service kiosk to check in before a flight, typed keywords into a search bar and received suggested results, or communicated with a digital assistant, you've interacted with AI.[2]

AI Technologies Include:

- Autonomous vehicles
- Chatbots
- Decision management
- Deep learning platforms
- Image generators
- Digital assistants
- Digital image processing
- Entertainment streaming apps
- Facial, speech, and image recognition
- Gaming
- Generative AI tools like ChatGPT
- GPS navigation
- Pattern recognition
- Fraud detection
- Virtual assistants
- Speech recognition
- Weather prediction
- Security and surveillance

3.2 History of Artificial Intelligence (AI)

The modern groundwork for AI began in the early 1900s, but the biggest strides were made in the middle of the 20th century, when pioneers like Alan Turing began exploring foundational concepts like artificial neural networks, machine learning, and symbolic reasoning.

The term artificial intelligence was coined and came into popular use in the mid-1950s following Turing's publication of "Computer Machinery and Intelligence," a paper that proposed a test of machine intelligence called the Imitation Game. Turing's publication eventually became the **Turing Test,** which experts used to measure computer intelligence. AI technology continued to develop from the mid-1950s to the late 1970s, as computers became faster, cheaper, more accessible, and could store more information. During this time, the first AI programming languages were created, machine learning algorithms were improved, and books and films began to explore the idea of robots. But computers were still millions of times too weak to exhibit intelligence. Research funding declined until the 1980s.

The 1980s were a period of rapid growth and interest in AI following breakthroughs in research and additional government funding. Deep learning techniques and the use of expert systems became more popular, both of which allowed computers to learn from their mistakes and make independent decisions.

The early 1990s showed some strides forward in AI research, including the first AI system that could defeat a reigning world champion chess player. The early 2000s saw innovations such as the first robot vacuum and the first commercially available speech recognition software.

The late 1990s and the early 2000s also saw significant advances in artificial intelligence. Computerized automation began to emerge, and machine learning was applied to many problems in academia and industry due to the availability of powerful computer hardware, immense collections of data, and the application of advanced mathematical methods.[3]

3.3 How Is Generative AI Different from Traditional AI?

Generative AI differs from traditional AI based on capabilities and applications. **Traditional AI** can mainly analyze data and make predictions. It excels at pattern recognition, so it can see data and then tell you what it sees. **Generative AI** can create new data and content informed by its training data. It excels at pattern creation, so it can see and tell you what they see in the data, but also use that data to create something new.

In other words, generative AI creates new content based on patterns in data and traditional AI focuses on analyzing and classifying existing data.[4]

3.4 Why is AI important?

AI is important because, for the first time, traditionally human capabilities can be undertaken in software inexpensively and at scale. AI can be applied to every sector to enable new possibilities and efficiencies.

AI technology is important because it enables human capabilities – understanding, reasoning, planning, communication and perception – to be undertaken by software increasingly effectively, efficiently and at low cost.

General analytical tasks, including finding patterns in data, that have been performed by software for many years can also be performed more effectively using AI. The automation of these abilities creates new opportunities in most business sectors and consumer applications.

Significant new products, services and capabilities enabled by AI include autonomous vehicles, automated medical diagnosis, voice input for human-computer interaction, intelligent agents, automated data synthesis and enhanced decision-making. AI has numerous, tangible use cases today that are enabling corporate revenue growth and cost savings in existing sectors.

Applications will be most numerous in sectors in which a large proportion of time is spent collecting and synthesizing data: financial services, retail and trade, professional services, manufacturing and healthcare. Applications of AI-powered computer vision will be particularly significant in the transport sector.

Use cases are proliferating as AI's potential is understood. We describe 31 core use cases across eight sectors: asset management, healthcare, insurance, law & compliance, manufacturing, retail, transport and utilities. We illustrate how AI can be applied to multiple processes within a business function (human resources).[5]

3.5 Application of AI In Agriculture

Artificial Intelligence has emerged as an essential field of agriculture of the 21st Century. It provided us with the support and solutions for increasing productivity, efficiency and sustainability in agriculture by the aid of large agricultural data analysis aimed to assist farmers make the right decisions at the right time. The key technologies are summarized here (especially for the cucumber culture and detection of leaf diseases).

Top Applications of AI in Agriculture:

- **Precision Farming:** AI uses satellite data, drones, and sensors to monitor crop health, soil composition, and environmental factors in real-time. This allows for site-specific, targeted application of fertilizers and pesticides, reducing waste.
- **Smart Irrigation and Water Management:** Algorithms analyze soil moisture levels and weather forecasts to automate irrigation, ensuring crops receive optimal moisture while reducing water consumption.
- **Crop and Soil Health Monitoring:** Computer vision tools can detect pests, weeds, and diseases by scanning crops. AI analyzes soil samples for nutrient deficiencies and pH levels, recommending tailored fertilizer application.
- **Robotics and Automated Harvesting:** AI-powered machines use cameras to identify and select ripe produce, reducing manual labor costs and increasing harvesting speed and precision.
- **Predictive Analytics for Yield Forecasting:** By analyzing historical data, weather patterns, and current crop conditions, AI assists farmers in predicting harvest yields and planning logistics.

- **Livestock Management:** AI systems monitor livestock health, tracking movement and behavior to detect illness or injury, ultimately improving animal welfare and farm productivity. [6]
- **Pest Detection and Control:** Models may observe pests and recommend suitable methods for control. We do not need to pollute the environment and damage crops by excess use of pesticides.

3.6 Challenges of AI in Agriculture

The primary challenges of AI in agriculture include high implementation costs, lack of quality data, poor rural connectivity, and low digital literacy among farmers. These barriers prevent widespread adoption, further compounded by technical limitations in adapting models to diverse field conditions, data privacy concerns, and lack of trust in automated, "black box" systems. [7]

Although developing the Artificial Intelligence (AI) technology is highly promising in a myriad of ways, especially for the agricultural application (for instance automated cucumber leaf disease classification), there are several problems of developing the AI which hinders the achievement of it and the realization of the benefits. These problems involve several problems related to technical, field deployment, socio-economic, ethical issues and the problems related to the domain of plant pathology. The computer vision application for image-based cucumber disease classification at the scope of this thesis has been systematically and critically analyzed on the grounds of evidence through this chapter.

Challenges of AI Adoption in Agriculture

1. **High Financial Costs and Unclear ROI:** The significant investment required for hardware, software, and sensors are a major barrier for many farmers, especially when the return on investment (ROI) is not immediately obvious.
2. **Data Quality and Availability:** AI requires vast amounts of high-quality data to make accurate predictions. In agriculture, data is often scattered, unstructured, or unavailable, which hinders the training of robust models.
3. **Infrastructure Limitations:** Lack of reliable internet connectivity and electricity in rural areas limits the functionality of cloud-based AI tools and IoT devices.
4. **Digital Literacy and Technical Skill Gaps:** Farmers often lack the necessary technical skills to effectively use, maintain, and interpret data from AI systems, leading to slow adoption.
5. **Model Adaptability and Generalization:** AI models trained in one environment (e.g., a specific farm type or region) often struggle to perform accurately in another due to variations in soil, climate, and crop types.
6. **Trust and Explainability:** Many AI systems act as "black boxes," making it difficult for farmers to understand how a decision was reached. This lack of transparency leads to distrust.

7. **Data Privacy and Security:** Concerns regarding who owns and controls the data generated on farms (e.g., satellite images, yield data) create anxiety over data misuse or corporate ownership.
8. **Environmental Variability:** Agricultural processes are subject to unpredictable weather and ecological changes, making it difficult for AI to provide consistent results compared to controlled environments. [8]

3.7 Why we should use AI in Agriculture

Artificial Intelligence (AI) in agriculture is transforming farming from a traditional, intuition-based practice into a data-driven, precision-focused industry. As the world's population increases, putting pressure on food systems, AI offers solutions to maximize yields, optimize resource use, and enhance sustainability

Here are the key reasons why we should use AI in agriculture,

1. Increased Crop Yields and Productivity

AI algorithms analyze data from soil sensors, drones, and satellites to provide real-time insights into plant health, nutrient deficiencies, and optimal harvesting times.
Precision Farming: AI ensures plants receive tailored care, boosting both the quantity and quality of produce.
Yield Prediction: AI models analyze historical data, weather trends, and crop performance to accurately forecast yield, allowing for better harvest logistics.

2. Effective Resource Management and Cost Reduction

AI helps farmers move away from blanket applications of fertilizers and pesticides, which cuts down on waste and expenses.
Precision Spraying: AI-enabled drones and robots can identify weeds or pests and apply herbicides or pesticides only where necessary, reducing chemical use by up to 90% in some cases.
Smart Irrigation: AI-driven systems process soil moisture data and weather forecasts to deliver the precise amount of water needed, reducing water consumption by 30–50%

3. Early Pest and Disease Detection

AI-powered computer vision can scan crop images (from drones or smartphones) to identify signs of stress, disease, or pest infestations faster and more accurately than human scouting.
Proactive Action: Early detection allows farmers to act before a problem becomes widespread, preventing significant crop losses.

4. Sustainability and Environmental Protection

Using AI contributes directly to environmental goals by reducing the ecological impact of farming.
Lower Emissions: Reduced use of fuel, chemicals, and water lowers the overall carbon footprint of farms.

Soil Health: Optimized inputs prevent over-application of chemicals, protecting soil quality.

5. Livestock and Dairy Health Monitoring

AI extends beyond crop management into livestock farming, enhancing animal welfare and productivity.

Real-time Monitoring: Wearable sensors and cameras track the health, behavior, feeding habits, and temperature of animals.

Early Warning Systems: AI can detect atypical behavior (like signs of illness or birthing), allowing farmers to react promptly to improve animal well-being.

6. Overcoming Labor Shortages

Autonomous machinery, such as robotic harvesters and self-driving tractors, fills the gap caused by the declining availability of farm workers.

Continuous Operation: Autonomous machines can work long hours, speeding up processes like planting and harvesting.

7. Climate Adaptation

As climate change brings unpredictable weather patterns, AI provides actionable insights for adaptation.

Better Forecasting: AI offers localized, precise weather forecasting that is more useful than traditional 14-day forecasts, helping farmers select suitable crops and plan the best times for planting.

3.8 How AI can improve the farming of cucumber plants in the field

AI can significantly improve field cucumber farming through precision agriculture, automated monitoring, and predictive analytics, which together boost yields, reduce resource wastage, and enable early disease detection. Key applications include using computer vision to detect leaf diseases with up to 96.81% accuracy, optimizing irrigation via soil sensors, and implementing automated harvesting robots. [9]

Here is how AI improves cucumber farming:

1. Precision Disease and Pest Management

Early Detection: AI-powered imaging systems (using drones or ground cameras) can identify pests, nutrient deficiencies, and diseases like downy mildew early, enabling targeted intervention before widespread crop loss.

Targeted Spraying: Rather than covering the entire field, AI-enabled systems can identify specific infected cucumber plants and apply fungicides or pesticides only where needed, reducing chemical use.

Image Classification: Advanced models, such as those using CNNs (Convolutional Neural Networks) and MobileNetV2, can classify cucumber leaf diseases with over 96% accuracy, allowing for fast, precise treatments. [10]

2. Optimized Irrigation and Resource Management

Smart Irrigation: AI integrates data from soil sensors, weather forecasts, and crop growth stages to determine the precise amount of water needed, preventing water waste and reducing drought risk.

Fertilizer Optimization: Machine learning algorithms evaluate soil nutrient levels in real time to recommend precise fertilizer application, minimizing environmental runoff and cutting costs. [11]

3. Automated Harvesting and Sorting

Autonomous Harvesters: AI-enabled robots, utilizing computer vision, can identify ripe cucumbers and harvest them without damaging the plant, addressing labor shortages and reducing manual labor costs.

Smart Sorting: AI systems can sort harvested cucumbers by size, color, and shape (e.g., differentiating between 11 types of fruit shapes), improving marketability. [12]

4. Yield Prediction and Growth Analysis

Predictive Analytics: By analyzing historical data, weather patterns, and soil conditions, AI models can accurately forecast yields, allowing farmers to adjust their strategies for better economic returns.

Growth Monitoring: Computer vision can monitor the shape, length, and tumor density of cucumbers, providing data-based insights to optimize phenotypes and improve quality. [13]

5. Environmental Control (For Protected Field/Tunnel Farming)

Autonomous Cultivation: AI algorithms can manage climate factors such as temperature, humidity, and CO_2 levels in covered field settings to optimize growth, with some studies showing potential for significant yield increases. [14]

CHAPTER 4
MACHINE LEARNING APPROACHES FOR PLANT DISEASE DETECTION

4.1 Introduction
In the last few years, machine learning (ML) and artificial intelligence have seen a new wave of publicity fueled by the huge and ever-increasing amount of data and computational power as well as the discovery of improved learning algorithms. However, the idea of a computer learning some abstract concept from data and applying them to yet unseen situations is not new and has been around at least since the 1950s. Many of these basic principles are very familiar to the pharmacometrics and clinical pharmacology community. In this paper, we want to introduce the foundational ideas of ML to this community such that readers obtain the essential tools they need to understand publications on the topic. Although we will not go into the very details and theoretical background, we aim to point readers to relevant literature and put applications of ML in molecular biology as well as the fields of pharmacometrics and clinical pharmacology into perspective. [15]

Researchers now use Machine Learning (ML) as their primary method for the automatic identification of plant diseases because this technology enables scientists to process agricultural diagnostic work without needing expert visual assessments. Applying ML techniques to cucumber (Cucumis sativus) leaf diseases demands that ML systems address three challenges, which encompass intricate visual patterns and two other factors.

For cucumber plants, machine learning approaches specialize in detecting common diseases like Downy Mildew, Powdery Mildew, Anthracnose, and Bacterial Wilt. Recent state-of-the-art models focus on Transfer Learning and lightweight Object Detection to handle complex field backgrounds and small lesions.[16]

Machine Learning (ML) techniques are widely used for automatic plant disease detection using image data. These approaches enable systems to learn patterns from cucumber leaf images and classify them into healthy or diseased categories. ML-based methods can be broadly divided into traditional approaches and deep learning approaches.
Machine Learning (ML) has become an essential tool in modern agriculture for solving complex problems such as plant disease detection. The system learns from data by discovering patterns which allow it to make decisions without needing explicit programming. The application of ML techniques to cucumber farming enables accurate and efficient disease detection through leaf image analysis.

The standard disease detection process requires manual inspection work, which takes a lot of time and leads to mistakes by human operators. The automated disease detection system from machine learning provides better accuracy with less required work because it enables doctors to discover diseases at earlier stages. Machine learning models use image data to identify minor changes in leaf color and texture and shape that humans cannot detect.

4.2 Traditional Machine Learning Approaches

Traditional ML methods involve two main steps: feature extraction and classification.

a) Feature Extraction

Important features are manually extracted from leaf images, such as:

- Color features (RGB, HSV values)
- Texture features (using GLCM – Gray Level Co-occurrence Matrix)
- Shape features (edges, contours)

b) Classification Techniques

After feature extraction, classifiers are used to identify diseases:

- Support Vector Machine (SVM) – Effective for binary and multi-class classification
- k-Nearest Neighbor (KNN) – Classifies based on similarity with nearest samples
- Decision Tree – Uses rule-based classification
- Random Forest – Combines multiple decision trees for better accuracy

Limitations:

- Requires manual feature design
- Sensitive to lighting and background variations
- Lower accuracy compared to deep learning

4.3 Application of Machine Learning

Machine learning (ML) and deep learning (DL) are revolutionizing agriculture by enabling fast, accurate, and automated detection of plant diseases through imaging, reducing economic losses, and enabling early treatment. Key applications include using Convolutional Neural Networks (CNNs) and models like YOLOv8 to identify pests and diseases, smartphone apps for real-time field diagnosis, and drone-based monitoring for large-scale surveillance. [17]

Key Applications of Machine Learning in Plant Disease Detection:

Leaf Disease Identification via Image Classification: Deep Learning models, particularly CNNs, analyze leaf images to classify healthy plants from infected ones, identifying specific diseases.

Real-time Mobile Applications: Smartphone-based tools allow farmers to take photos of plant leaves and receive immediate diagnosis and treatment recommendations, facilitating quick, early-stage intervention.

Automated Disease Diagnosis with Object Detection: Advanced algorithms like Faster R-CNN and SSD (Single Shot Multibox Detector) are used to locate and identify multiple pests and diseases on plant parts.

Drone and Aerial Imagery Monitoring: ML algorithms process aerial images from drones and satellites, allowing for efficient, automated scanning of vast agricultural fields.

Smart Farming and Internet of Things (IoT): AI is integrated with IoT sensors for continuous monitoring, helping to predict diseases by analyzing environmental data alongside visual symptoms.

Genomic Analysis for Disease Resistance: Machine learning helps analyze genetic data to develop resistant crops, allowing farmers to select optimal genomic sequences to combat pathogens.[18]

Commonly Used Techniques:

- **Convolutional Neural Networks (CNNs):** Often considered the most effective for image-based identification.
- **Support Vector Machines (SVM) & Random Forests:** Traditional supervised learning methods used for classification.
- **Image Processing & Feature Extraction:** Techniques like GLCM (Grey Level Co-occurrence Matrix) are used for texture classification to detect early disease symptoms.[19]

4.4 Challenges of Machine Learning

Machine learning (ML) in plant disease detection faces significant challenges, including limited, imbalanced datasets, poor generalization to real-world field conditions, and low computational efficiency on edge devices. While models achieve high accuracy in controlled lab settings, they often struggle with noisy environments, varying illuminations, and diverse plant species, reducing performance significantly in practical applications.[20]

4.4.1 Key challenges include:

- **Data Scarcity and Quality:** Insufficient labeled images for certain diseases and unbalanced datasets (too many healthy, too few diseased) hinder model training.
- **Real-World Environmental Noise:** Factors like complex backgrounds, changing illumination, and weather conditions reduce accuracy compared to laboratory environments.
- **Model Generalization:** Models trained on a specific dataset often fail when applied to different environments, seasons, or camera types, limiting their scalability.
- **Symptom Variability:** Early-stage diseases are hard to detect, and similarities between different diseases can cause misdiagnosis.
- **Computational Constraints:** Implementing complex deep learning models on resource-limited devices (smartphones, IoT sensors) in the field requires balancing accuracy with speed.[21]

4.4.2 Model & Algorithmic Limitations

- **Accuracy-Latency Trade-off:** High-capacity models (ViT, ResNet101) exceed memory/compute limits on agricultural edge devices, while lightweight architectures (MobileNetV3) sacrifice 2–4% accuracy.
- **Symptom Mimicry:** Nutrient deficiencies, herbicide injury, and pest damage visually resemble biotic diseases, causing false positives in single-label classifiers.
- **Black-Box Opacity:** Complex ensembles and deep networks lack transparent decision pathways, reducing farmer trust and hindering regulatory acceptance.
- **Calibration Drift:** Overconfident predictions on out-of-distribution samples (new cultivars, seasonal variations) lead to inappropriate treatment recommendations.

4.4.3 Cucumber-Specific Pathological Challenges

Rapid Disease Progression: Downy Mildew and Angular Leaf Spot can spread visibly within 24–48 hours under high humidity, demanding high-frequency monitoring and low-latency inference.[22]

Green-on-Green Contrast: Early chlorotic lesions exhibit low color differentiation against healthy cucumber tissue, requiring high-resolution capture and texture-aware architectures.[23]

Cultivar & Stage Variability: Symptom expression differs across hybrid vs. open-pollinated varieties and growth stages, necessitating domain-adaptive fine-tuning or stratified evaluation protocols.[24]

CHAPTER 5
DEEP LEARNING ARCHITECTURES FOR CUCUMBER LEAF DISEASE DETECTION

5.1 Introduction

Agriculture is a vital pillar of every economy, significantly influencing a country's economic development [25,26]. It not only generates essential raw materials and food products but also creates numerous job opportunities [27]. However, agriculture faces myriad challenges, such as irregular rainfall and poor soil fertility [28]. Among these challenges, one issue stands out prominently: the infestation of crops by diseases [29]. Plant diseases significantly impact global agricultural productivity, leading to yield losses and economic damage [30,31]. Cucumbers are one of the most important greenhouse crops in Europe, as well as in the Near and Far East [32]. Global cucumber production amounted to 88 million tons in 2019, 80% of which is produced in China [33]. A large proportion of production also takes place in Iran, Turkey, the USA, Japan, and Russia [32]. In these and many other countries, cucumbers are an important part of society's nutrition [34]. They are particularly susceptible to various diseases caused by adverse environmental conditions [30]. These diseases, including Anthracnose, Bacterial Wilt and Downy Mildew, can cause considerable harm and create gaps in food supply chains [30,35]. Downy Mildew, for example, can lead to yield losses up to 80%, as well as seed losses of up to 70% [36]. Early detection and treatment are crucial to improve crop yields and support farmers, especially in regions like Bangladesh, where the used dataset was collected [37]. By diagnosing the disease at an early stage, which is essential as fruit and vegetables spoil quickly, farmers can reduce their use of pesticides in the long term to increase the quality and sustainability of their harvest [38].

For this reason, there is an urgent need for efficient and accurate disease detection methods, given the importance of early diagnosis for plant diseases [39]. Traditional methods are often time-consuming and inaccurate, highlighting the increasing focus on automated solutions like machine learning and deep learning technologies [30,35]. Human diagnosis of diseases is prone to errors and can lead to diseases going undetected [40]. This can endanger the entire crop. Machine learning models are faster and more precise [40]. They detect diseases earlier, which allows farmers to respond in a timely and appropriate manner [41].

Existing plant disease detection technologies, for example, use a MobileNet V3 network that is specially optimized for mobile devices [42]. However, similar procedures, although they achieve good results, are still underdeveloped and often expensive [42]. Plant photographs for disease detection are frequently of poor quality due to close planting [43]. Traditional methods based on human inspection are inefficient, particularly in the early stages of disease [42,28]. Developing a machine learning model based on a specialized dataset that displays cucumber diseases in a hyperspectral array of images is therefore crucial [44]. Many recognition and diagnostic methods follow the pipeline method of image segmentation, feature extraction and pattern recognition. However, two problems arise when using the pipeline method: First, the accuracy of such methods is highly dependent on how the visible disease features are selected and extracted. Moreover, the complexity of these approaches is notably increased due to the variable lighting conditions often encountered in field images [45]. Research has also shown that the background of images significantly influences detection and that identifying diseases from a single lesion in an image can alleviate the problem of data scarcity for training deep learning models [46].

Key Aspects of Deep Learning for Cucumber Disease Detection:

- ❖ **Core Technologies:** CNN architectures such as VGG16, VGG19, ResNet50, InceptionV3, and YOLO (You Only Look Once) are commonly employed for feature extraction and classification.
- ❖ **Targeted Diseases:** Models are trained to detect various diseases, including Anthracnose, Bacterial Wilt, Downy Mildew, Belly Rot, and Gummy Stem Blight, ensuring healthier crops.
- ❖ **Improved Accuracy:** Transfer learning (using pre-trained models) combined with fine-tuning techniques, such as freezing layers and adjusting input resolutions (e.g., 224x224), significantly increases accuracy to over 98%.
- ❖ **Real-time Applications:** Lightweight models like YOLO-Cucumber are deployed in mobile or embedded systems, enabling real-time detection in agricultural fields.
- ❖ **Overcoming Challenges:** Techniques such as data augmentation, segmentation, and specialized loss functions (e.g., Target-aware Loss) address data scarcity and improve detection of small or early-stage lesions. [47]

These advanced systems directly enhance crop productivity and sustainability by enabling precise, early detection of plant stress.[48]

5.2 Convolutional Neural Network (CNN)

In recent years, as state-of-art machine learning, deep learning-based algorithms have been widely used in many different fields in terms of recognition and classification. Deep learning as a new area of machine learning has become well known in plant disease recognition. It is possible to offer promising results and be efficient for learning and processing from the composite input data, which includes feature extraction and classification at the same time. The convolutional neural network (CNN) is one of the best-performing techniques of deep learning algorithms for image recognition in the machine learning field. This technique automatically acquires requisite features from the input image during the learning process which is the most important advantage of deep learning compared with traditional machine learning algorithms [9]. Despite reducing the complicated handcraft engineering processes, CNN is used to obtain a good performance of classification accuracy [10]. CNN has a good performance in terms of cucumber disease diagnosis, general image categorization, and leaf classification [11], as well as in flower recognition [12]. In [9], we realized that deep CNN in plant disease recognition is remarkably higher than the traditional algorithms. Additionally, in [13], the authors found that deep learning performed better than other machine learning methods to classify normal and diseased cells. In [14], long short-term memory (LSTM) is also used as a classification method to gear fault diagnosis.

However, none of the existing literature has considered ResNet-50, Inception-V3, and two combined CNN architectures for cucumber leaf disease recognition. In other words, none of the studies have worked on spider cucumber disease types that are related to pest types. On the other hand, CNN has a drawback in terms of using a huge amount of data, whereas available datasets do not have sufficient image data needed to do the proper recognition [15]. In the area of cucumber disease recognition, there is a lack of adequate standard public datasets for research purposes, and this considerably affects the efficiency of the classification process. To address those issues, two different cucumber disease datasets have been used to compute the experimental results: the locally constructed dataset and the cucumber plant diseases achieved online which are publicly available. In such circumstances, cucumber disease diagnosis in terms of accuracy, low cost, and speed are in great demand, and still, it is challenging for the agricultural industry and farmers.[49]

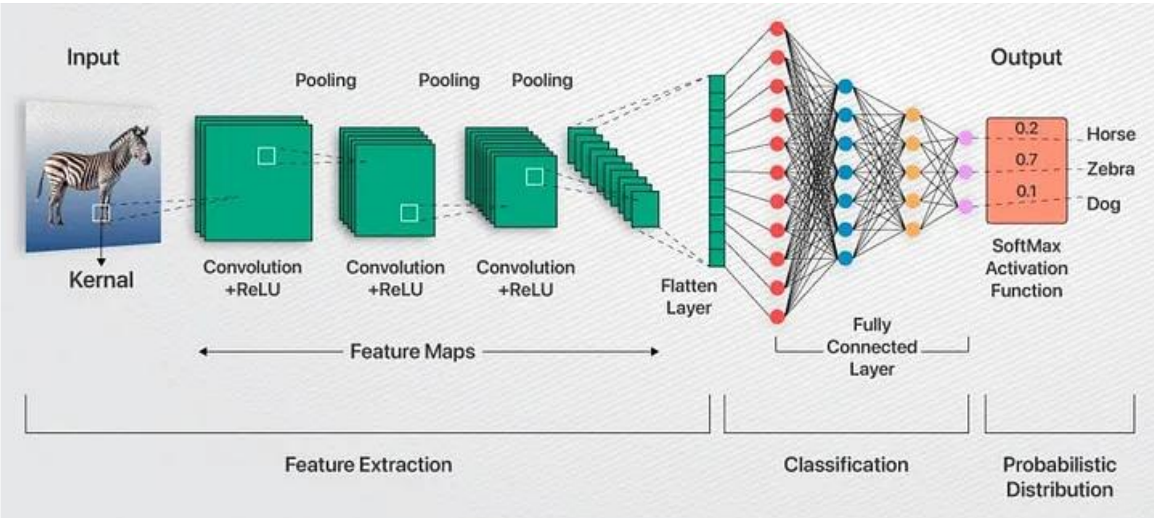

Fig 5.1: CNN Architecture

Key Components:

- **Convolutional Layers:** Extract features from images
- **Activation Function (ReLU):** Introduces non-linearity
- **Pooling Layers:** Reduce image size and complexity
- **Fully Connected Layers:** Perform classification

CNNs are widely used due to their high accuracy and ability to learn spatial features from images.

Key Techniques and Architectures

Common Architectures: Studies frequently use VGG16, Inception-V3, ResNet50, and DenseNet201 for high accuracy, sometimes exceeding 99% in validation.

Custom/Advanced CNNs: A Global Pooling Dilated CNN (GPDCNN) can replace fully connected layers to reduce parameters and use dilated kernels to enlarge the receptive field for better feature extraction.

Data Preprocessing: Data augmentation (rotation, translation) is crucial to prevent overfitting and improve robustness.

Segmentation: Utilizing techniques like DeepLabV3+ and U-Net helps isolate disease spots from complex backgrounds, achieving higher accuracy in identifying disease severity.[50]

5.3 MobileNetV2

MobileNetV2 is a highly effective convolutional neural network (CNN) for cucumber leaf disease detection, particularly when deployed on mobile devices or low-resource platforms. Its lightweight architecture uses inverted residual blocks and depth wise separable convolutions, which significantly reduce the number of parameters and computational costs while maintaining high accuracy for real-time field diagnosis.[51]

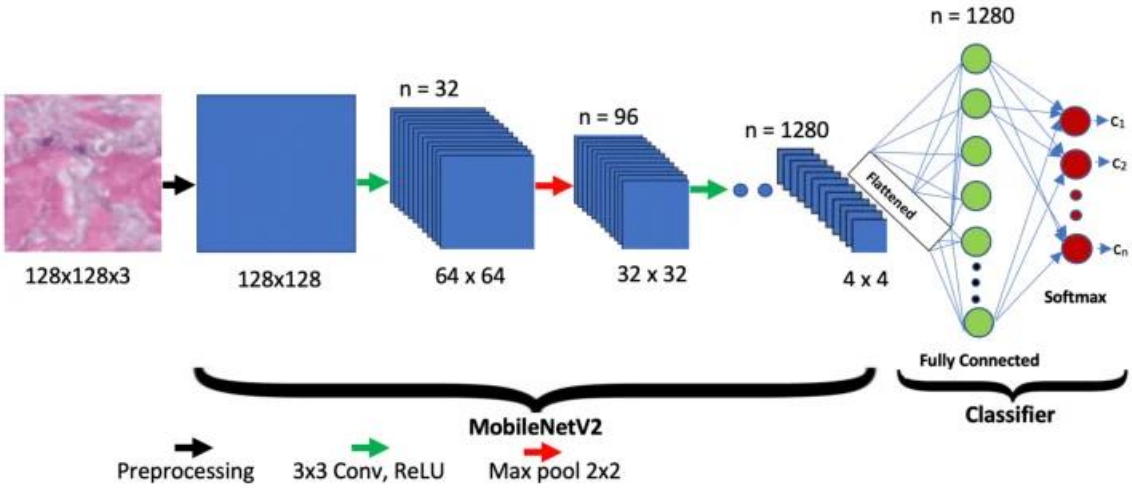

Fig 5.2: Mobilenetv2 Architecture

5.3.1 Key Innovations in MobileNetV2

- ❖ **Depth wise Separable Convolutions:** Decomposes standard convolutions into depth wise (channel-wise filtering) and pointwise (1×1 cross-channel mixing) operations, reducing FLOPs by 8–9× compared to standard CNNs.
- ❖ **Inverted Residual Blocks:** Expands low-dimensional inputs to high-dimensional feature spaces using 1×1 convolutions, applies depth wise filtering, then compresses back via linear bottlenecks. This preserves information flow while minimizing parameter overhead.
- ❖ **Linear Bottlenecks:** Removes non-linear activation (ReLU6) in the final compression layer to prevent information loss in low-dimensional feature maps, critical for preserving fine-grained disease textures.
- ❖ **Skip Connections:** Residual pathways enable gradient flow across deep networks, mitigating vanishing gradients during fine-tuning on limited agricultural datasets.

5.3.2 Why MobileNetV2 for Cucumber Leaf Disease Detection?

MobileNetV2 is strategically selected for real-time, field-deployable cucumber disease scanning. Smallholder farmers and extension workers typically operate mid-range smartphones with constrained RAM and processing power. MobileNetV2's ~3.4M parameter footprint and <200 MB RAM requirement enables offline inference with latency under 400 ms per image. Despite its lightweight design, it retains sufficient representational capacity to distinguish cucumber foliar diseases, particularly when fine-tuned with domain-aware augmentation. Its efficiency supports continuous field monitoring without draining device batteries or requiring cloud connectivity, aligning directly with rural agricultural constraints.

5.4 ResNet50

ResNet-50 is a 50-layer Convolutional Neural Network (CNN) designed for image recognition, winning the 2015 ILSVRC task with a ~3.57% error rate. It utilizes residual learning, or "skip connections," to overcome the vanishing gradient problem, enabling deeper network training. It is widely used in computer vision for classification.

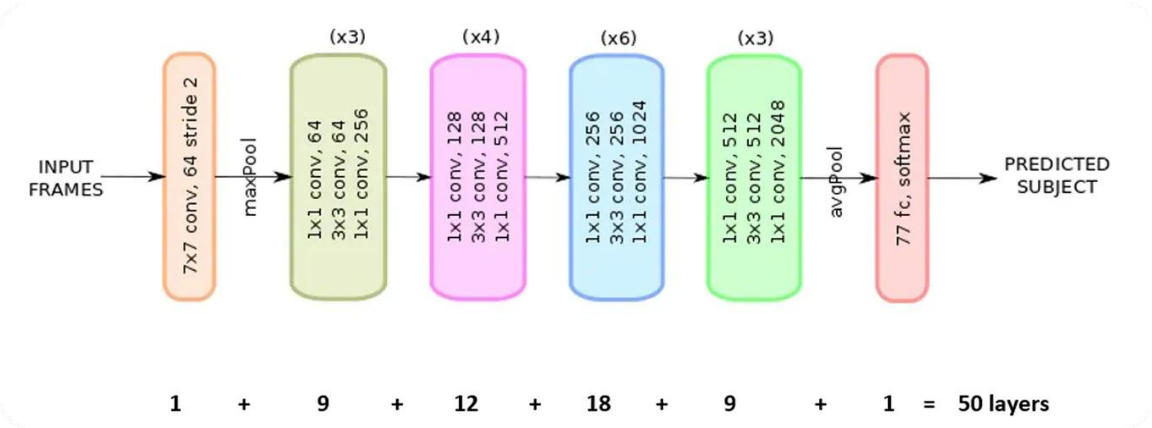

Fig 5.3: ResNet50 Architecture

5.4.1 Key Features of ResNet50

The introduction of residual connections which enable gradients to bypass intermediate layers through skip connections, which direct backpropagated gradients to their final destination inside the architecture, created a fundamental transformation that ResNet50 brought to deep learning. The complete system consists of:

The system contains 50 weighted layers, which are divided into four primary stages that each use increasing filter depths of 64, 128, 256 and 512.

The bottleneck blocks use 1x1 to 3x3 to 1x1 convolutional pattern to achieve optimal computational efficiency while extending their receptive field capability

The system uses global average pooling together with a fully connected classifier head to process its input data.

The residual design method successfully prevents vanishing gradients, which enables deep networks to maintain stable training while achieving exceptional performance in extracting essential multi-scale features that are needed to interpret complex agricultural images.

5.4.2 Why ResNet50 for Cucumber Leaf Disease Detection?

The thesis establishes ResNet50 as the baseline method which achieves the highest accuracy for its experiments. The system uses deep hierarchical feature extraction to identify visually similar cucumber diseases, which include early Downy Mildew and Angular Leaf Spot. The architectural design is proven to deliver dependable results, which produces reliable outcomes in both testing and actual agricultural research studies conducted in real world conditions. ResNet50 requires more computational power than MobileNetV2, but it establishes an important performance standard to determine the highest possible classification accuracy before implementing edge performance improvements.

5.5 EfficientNet

EfficientNetB0 is a highly efficient, lightweight convolutional neural network designed for image classification, achieving state-of-the-art accuracy with only 5.3 million parameters. Part of the EfficientNet family, it uses compound scaling (optimally balancing depth, width, and resolution) to outperform larger models. It operates on 224x224 input images, often serving as a robust, fast backbone for transfer learning in computer vision tasks.

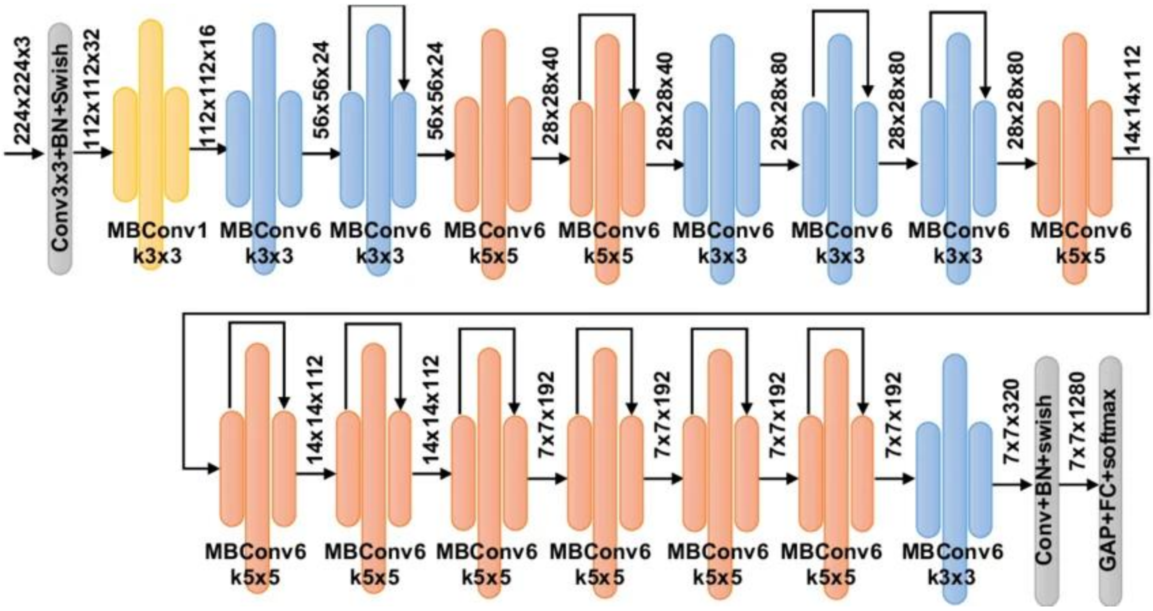

Fig 5.4: EfficientNet B0 Architecture

5.5.1 Key Features of EfficientNetB0

EfficientNetB0 uses compound scaling as its main approach which scales network depth and width and input resolution through fixed scaling coefficients that result from neural architecture search. Key components include:

MBConv (Mobile Inverted Bottleneck Convolution) blocks with squeeze-and-excitation (SE) attention to dynamically recalibrate channel-wise feature importance

Swish activation functions that improve gradient flow and model expressivity compared to ReLU

A carefully balanced parameter budget (~5.3M) that maximizes accuracy per FLOP

The implementation of compound scaling enables architectural growth throughout all dimensions which eliminates both bottlenecks and redundant components that would occur under arbitrary scaling methods.

5.5.2 Why EfficientNetB0 for Cucumber Leaf Disease Detection?

(Note: Thesis standardized to B0 for architectural consistency; B3/B4 variants can be deployed if server-side accuracy is prioritized.)

The research team selected EfficientNetB0 because it provides the best combination of accuracy and resource efficiency. The MBConv+SE architecture of EfficientNet enables cucumber disease classification through its ability to recognize multiple scale lesion patterns which need both detailed texture analysis and wide spatial understanding. The B0 model achieves better performance than both ResNet50 and MobileNetV2 according to empirical tests which measure results at the same or lower resource usage. Its simplified structure allows for easy quantization and pruning which makes it perfect for use in both cloud-based precision agriculture dashboards and field applications that require edge optimization.

5.6 Deep Learning Applications

Deep learning applications for cucumber leaf disease detection leverage convolutional neural networks (CNNs) to identify diseases like Downy Mildew, Powdery Mildew, and Anthracnose with over 97-99% accuracy. Techniques including transfer learning (VGG16, ResNet50), YOLO-based real-time detection, and image segmentation enable high-precision diagnosis.[52]

Key Applications and Technologies

- **High-Accuracy Classification:** Models such as ResNet50V2 and VGG16 have achieved accuracies of 99.37% and 98.76% respectively, for identifying healthy vs. diseased leaves.
- **Real-Time Detection:** YOLO-based models (e.g., YOLO-Cucumber, YOLOv5) are implemented on embedded devices like NVIDIA Jetson for real-time monitoring, achieving upwards of 23 frames per second (FPS).
- **Advanced Architectures:** Researchers use specialized models like EfficientNet for high precision and hybrid frameworks (e.g., U-Net for segmentation combined with CNNs) to isolate lesions from complex backgrounds. [52]

Commonly Detected Diseases

- **Downy Mildew & Powdery Mildew:** Frequently classified using pre-trained VGG16, ResNet, and EfficientNet models.
- **Bacterial Wilt & Anthracnose:** Identified using VGG19 and custom CNN architectures.
- **Leaf Miner:** Detected by CNNs in high-precision studies.

5.7 Deep Learning Challenges

Deep learning for cucumber leaf disease detection faces significant challenges, including a lack of diverse, high-quality public datasets, the need for real-time inference on mobile devices, and high environmental variance in field conditions. Models often struggle with complex backgrounds, small, early-stage lesion detection, and differentiating between similar symptoms.[53]

Key deep learning challenges for cucumber leaf disease detection include:

- ❖ **Data Scarcity and Quality:** There is a lack of comprehensive, publicly available datasets containing diverse cucumber diseases, particularly in real-world, non-laboratory environments.
- ❖ **Complex Field Conditions:** Models struggle with varying lighting conditions, dirt, background interference, and low contrast images which reduce identification accuracy.
- ❖ **Early Disease Detection:** Detecting small lesions or subtle early-stage symptoms is difficult, yet crucial for timely, effective treatment.
- ❖ **In-Field Variability:** Factors such as overlapping leaves, pest damage, and multiple concurrent diseases complicate, and can reduce, detection accuracy.
- ❖ **Model Performance Trade-offs:** Balancing high classification accuracy with low computational cost for real-time application on devices like drones or smartphones (e.g., using Jetson Xavier) is a significant bottleneck.
- ❖ **Generalization:** Models trained on laboratory images often fail to generalize to new, unseen field data.[53]

5.8 Deep Learning Approaches

Deep learning approaches for cucumber leaf disease detection primarily utilize Convolutional Neural Networks (CNNs), achieving high accuracy (95-98%+) by identifying diseases like Downy Mildew, Anthracnose, and Bacterial Wilt. Top methods include transfer learning with VGG16/VGG19, ResNet50V2, and lightweight models like YOLOv5s for real-time edge computing. [54]

Key Deep Learning Approaches:

Transfer Learning & CNNs: Pre-trained models like VGG16, ResNet50, and InceptionV3 are widely used, with studies showing VGG16 achieving up to 98.76% accuracy, surpassing custom CNNs.

Lightweight Real-time Detection: YOLOv5s models, combined with image enhancement and data augmentation, are deployed on mobile/edge devices for immediate infield identification, achieving over 89% mean average precision.

Advanced Architectures: Innovative approaches include custom CucuNet-CNNs, Spiking Neural Networks (SNN) (reaching 98.91% accuracy), and dilated/inception CNNs for improved precision in field conditions.

Segmentation & Severity Analysis: Models like DeepLabV3+ and U-Net are used in tandem for two-stage segmentation (separating leaf from background, then identifying lesion spots) to classify disease severity.

Optimization Techniques: Techniques such as Deformable Convolutional Networks (DCN) and channel pruning are used to make detection models more efficient for early-stage disease identification and mobile deployment.[55]

CHAPTER 6
TOOLS & METHODOLOGY

6.1 Introduction
The chapter describes the process which led to our creation of an automated system that uses artificial intelligence to identify and categorize cucumber leaf diseases. The pipeline includes steps to obtain data and to prepare data and to select a model and to train the model and to implement the model for actual production use. Our system combines computer vision with learning technology to solve major agricultural challenges which include diagnostic delays and errors made by scouts and insufficient access to plant disease specialists. The system

The system is designed to provide reliable performance and strong protection and efficient operation. The system enables farmers to grow cucumbers by providing them with an artificial intelligence tool which accurately identifies plant diseases. The system enables users to operate the software under actual farming conditions. The solution offers cucumber farmers an efficient method which they can expand according to their needs. The developed system directly supports the objective of providing an AI-driven tool for cucumber cultivation. Our system provides an effective solution to the problem of cucumber leaf diseases.

This chapter presents the tools, dataset preparation, and methodology used for cucumber leaf disease detection and classification. The system developed by the researchers uses a custom Convolutional Neural Network (CNN) together with transfer learning models and machine learning algorithms to deliver precise results.

The models used in this research include:
- Custom CNN (4-layer architecture)
- MobileNetV2
- ResNet50
- EfficientNetB0
- SVM (Support Vector Machine)
- Random Forest

6.2 Data Collection
6.2.1 Data Sources
The training corpus was built with high-variance characteristics through image collection from various reliable sources:
- **Public Repositories:** Kaggle agricultural collections, and university extension archives.

- **Web-Scraped Agricultural Imagery:** Researchers collected this material from verified agronomy publications and research papers and open-access crop databases.

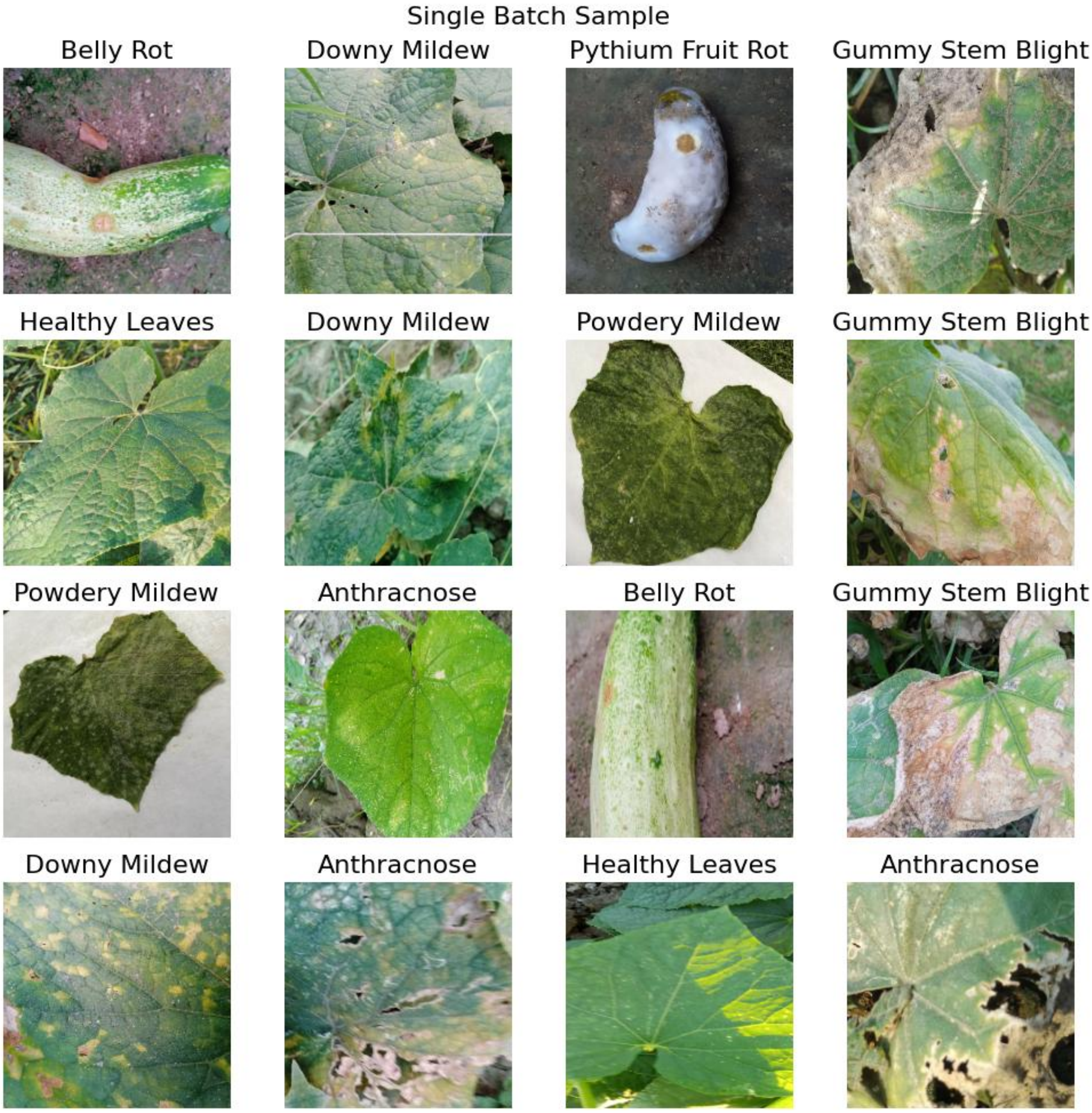

Fig 6.1: Dataset Sample

6.2.2 Hierarchical Deduplication Framework

We developed a rigorous deduplication process to create a unique collection of **X distinct images**. This was achieved through a two-stage filter using multiple hashing techniques:

- **MD5 Hash:** To catch exact file duplicates.
- **dHash, pHash, and aHash:** To find visually similar images that might have different file sizes or slight edits. This process eliminated the risk of data leakage between the training and testing sets, ensuring a fair and high-performing model.

6.2.3 Dataset Composition & Class Distribution

Class	Total Images	Training (70%)	Validating (15%)	Testing (15%)
Anthracnose	4510	3157	676	677
Bacterial Wilt	4231	2962	635	634
Belly Rot	3149	2204	472	473
Downy Mildew	4853	3397	728	728
Gummy Stem Blight	3149	2204	472	473
Healthy Leaves	6469	4528	970	971
Powdery Mildew	1989	1392	298	299
Pythium Fruit Rot	953	667	143	143
Total	29,303	20,511	4,394	4,398

Table 6.2.3 Dataset Composition & Class Distribution

6.3 Description of Cucumber Leaf Diseases

Accurate classification requires knowledge of the specific visual characteristics which different disease classes show because these characteristics guide the convolutional feature learning process.

1. Anthracnose

- Causal agent: Colletotrichum orbiculare (fungus)
- **Symptoms:**
  - Dark brown to black **sunken lesions** on leaves
  - Irregular spots with **yellow halos**
  - Can spread to stems and fruits
- **Favorable conditions:** Warm, humid weather
- **Impact:** Severe infection leads to leaf drop and reduced yield

  📷 Image: leaf with black sunken spots

2. Bacterial Wilt

- Causal agent: Erwinia tracheiphila (bacteria)
- **Symptoms:**
  - Sudden wilting of leaves without yellowing
  - Leaves may recover at night but worsen later
  - Sticky white fluid may appear when stem is cut
- **Transmission:** Spread by cucumber beetles

- **Impact:** Plant dies quickly if untreated

📷 Image*:* wilted cucumber plant

3. Belly Rot

- Causal agent: Soil-borne fungi (e.g., Rhizoctonia)
- **Symptoms:**
    - Water-soaked spots on fruit underside
    - Brown rot on lower leaves touching soil
- **Favorable conditions:** Wet soil, poor drainage
- **Impact:** Fruit becomes unmarketable

📷 Image: rotting cucumber fruit on soil

4. Downy Mildew

- Causal agent: Pseudoperonospora cubensis
- **Symptoms:**
    - Yellow angular spots on upper leaf surface
    - Gray/purple fuzzy growth underneath
- **Favorable conditions:** Cool, humid environment
- **Impact:** Rapid leaf death → reduced photosynthesis

📷 Image: yellow patches + underside mold

5. Gummy Stem Blight

- Causal agent: Didymella bryoniae
- **Symptoms:**
    - Brown lesions on stems and leaves
    - Sticky **gummy substance** oozing from stems
    - Leaf spots may enlarge and dry out
- **Impact:** Weakens plant structure and reduces yield

📷 Image: stem with gummy secretion

6. Healthy Leaves

- **Characteristics:**
    - Bright green color
    - No spots, lesions, or discoloration
    - Uniform texture and structure

- **Importance in dataset:**
    - Helps model distinguish **normal vs diseased**

📷 Image: healthy cucumber leaf

7. Powdery Mildew

- Causal agent: Erysiphe cichoracearum
- **Symptoms:**
    - White powder-like fungal coating on leaves
    - Leaves turn yellow and dry over time
- **Favorable conditions:** Dry but humid environments

- **Impact:** Reduces photosynthesis → poor growth

  📷 Image: white powder on leaf

8. Pythium Fruit Rot

- Causal agent: Pythium spp. (water mold)
- **Symptoms:**
  - Soft, watery rot on fruits
  - Leaves near soil may also decay
- **Favorable conditions:** Excess moisture
- **Impact:** Rapid fruit decay and crop loss

  📷 Image: soft rotten cucumber

6.4 Data Preprocessing

Data preprocessing forms an essential component of any deep learning model development process. The process ensures that input images maintain a clean and consistent appearance which makes them suitable for training purposes. This process leads to better system performance in accuracy and speed and increases model reliability.

Why Data Preprocessing is Important

Removes noise and inconsistencies

- Standardizes image format
- Helps model learn better patterns
- Prevents overfitting
- Improves generalization performance

6.4 1. Redundancy Check (Duplicate Removal)

The first step in the process involves checking for all duplicate elements which must be eliminated from the system. The purpose of this step creates the need to remove all matching or duplicate images from the system. The process holds significant value because duplicate data creates two major problems which first lead to model bias and second result in extra computations. The process results in a dataset which contains only distinct and uncorrupted data elements.

6.4.2 Resizing

To ensure compatibility with standard Convolutional Neural Network (CNN) architectures, all images were resized to **224×224 pixels**. A **center-crop padding** technique was applied during this process to maintain the original aspect ratios and prevent image distortion. This standardization ensures uniform tensor dimensions, allowing for efficient batch processing during model training.

6.4.3. Shuffling

The process randomly changes the arrangement of dataset images.

Why important:

- The method prevents the model from establishing sequential pattern recognition.

- The process creates balanced training conditions.
- The outcome leads to improved generalization ability.

6.4.4. Normalization

To improve computational efficiency, all pixel values were transformed from their original integer range of [0, 255] to a normalized floating-point range of [0, 1]. This scaling is achieved using the following formula:

$$\textbf{Normalized Pixel} = \frac{\text{Original Pixel Value}}{255}$$

The implementation of normalization provides three primary advantages for the deep learning model:

- **Accelerated Convergence:** It allows the training process to achieve faster results by ensuring the weight updates remain within a manageable range.
- **Numerical Stability:** It creates a more stable training environment, preventing large gradients from causing erratic behavior in the model.
- **Feature Balancing:** It ensures that all input features contribute equally to the learning process, preventing higher intensity pixels from disproportionately dominating the model's weight adjustments.

6.4. 5. Data Augmentation

The technique creates additional data resources to expand the existing dataset through its capacity to create new samples.

1. **Random Flip**

The tool performs horizontal and vertical image flips
The tool enables the model to recognize different image angles

2. **Random Rotation**

The tool performs image rotations through its random angle selection process
The process enables the model to handle different perspective changes

3. **Random Zoom**

The tool performs zoom operations through its random selection of zoom levels
The process enables the system to identify features which exist at multiple size dimensions

4. **Random Contrast**

The tool modifies both brightness and contrast settings
The process enables better performance when different lighting conditions exist

The primary advantage of the system

- The system prevents overfitting which leads to better actual performance results

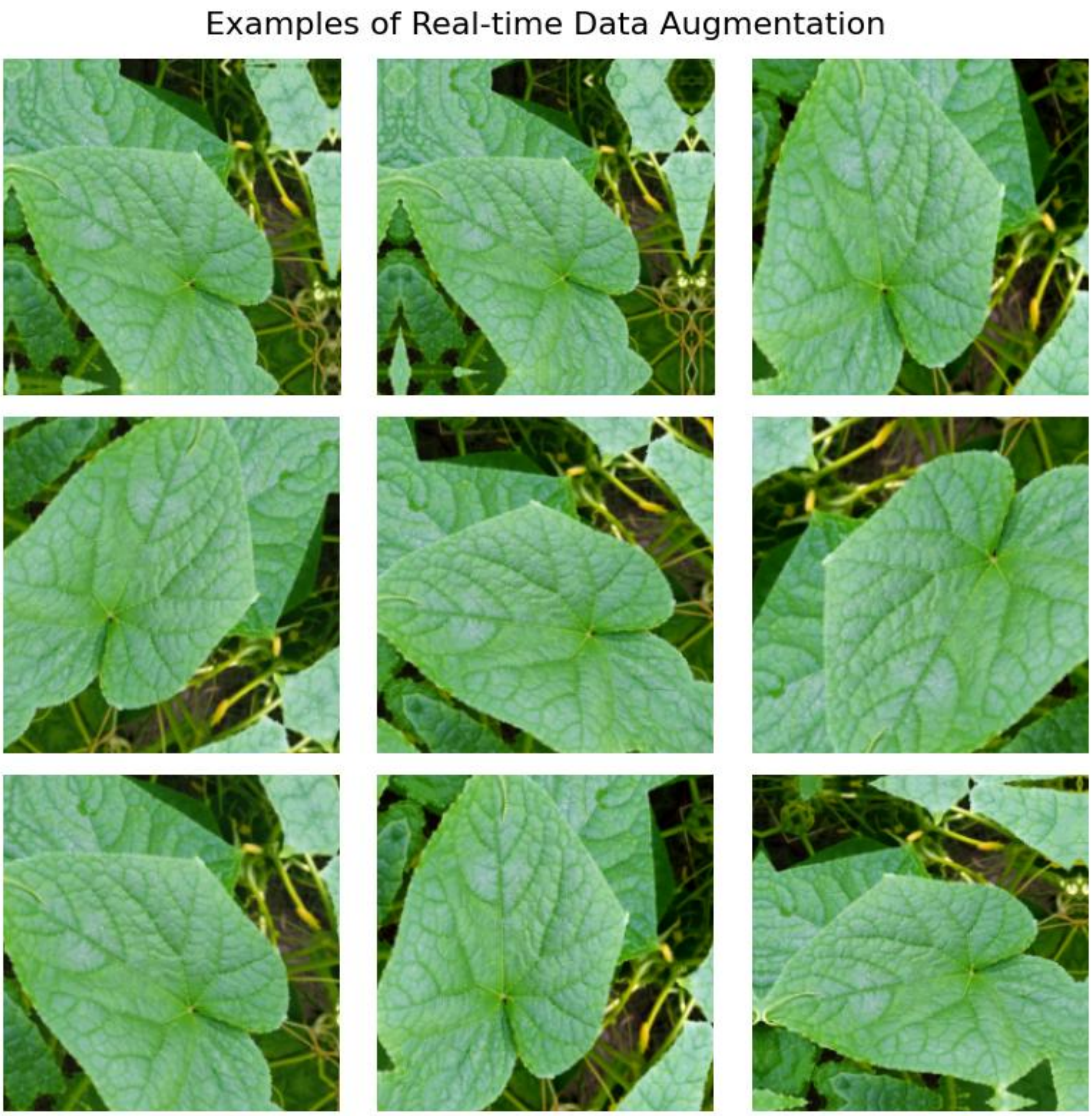

Fig 6.2: Data Augmentation Sample

6.4. 6. Autotune Optimization

- ❖ The system automatically optimizes the data loading process from start to finish
- ❖ The system operates in multiple frameworks including TensorFlow
- ❖ The system enables
- ❖ The system enables
- ❖ The system enables training processes to be completed in reduced timeframes.

6.5 Model Training & Evaluation

6.5.1 Model Training

The dataset is divided into three parts:

- **Training Set:** 70%
- **Validation Set:** 15%
- **Testing Set:** 15%

Training Configuration:

- **Loss Function:** Categorical Crossentropy
- **Optimizer:** Adam
- **Evaluation Metrics:** Accuracy, Loss, Precision, Recall

Training is performed iteratively to optimize model weights and improve classification performance.

6.5.2 Transfer Learning Approaches

In addition to the custom CNN model, transfer learning is applied using pre-trained models.

Models Used:

- **Custom CNN (4-layer)** – Trained from scratch
- **MobileNetV2** – Lightweight and efficient
- **ResNet50** – Deep architecture with residual connections
- **EfficientNetB0** – Balanced scaling of network parameters

Transfer Learning Strategy:

- Use pre-trained weights (e.g., ImageNet)
- Freeze initial layers

The last three models (**MobileNetV2, ResNet50, EfficientNetB0**) are implemented using transfer learning to improve performance and reduce training time.

6.5.3 Model Evaluation

Custom CNN (4-layer architecture)

The core of this research is a custom-designed, 4-layer Convolutional Neural Network (CNN) optimized for plant pathology classification. The model follows a sequential architecture designed to extract spatial features efficiently while maintaining computational speed.

The architecture consists of the following key components:

1. **Convolutional Layers:** The model features four specialized 2D convolutional layers.
    - The first layer uses **32 filters**, followed by layers with **64**, **128**, and **128 filters** respectively.
    - Each layer uses a 3×3 kernel to capture essential patterns such as leaf spots, wilting textures, and fungal growth.
2. **Downsampling (Pooling):** Following each convolution, a **MaxPooling2D** layer (2×2) is used. This reduces the spatial dimensions (from 224×224 down to 14×14), which helps the model focus on the most important features and reduces the risk of overfitting.
3. **Fully Connected Layers:**
    - A **Flatten** layer converts the 2D feature maps into a 1D vector of **25,088 elements**.
    - A **Dense (Hidden) Layer** with **256 neurons** performs the high-level reasoning.

- A **Dropout Layer (rate = 0.5)** is implemented to prevent the model from relying too heavily on specific neurons, ensuring better generalization to new leaf images.

4. **Output Layer:** The final Dense layer consists of **8 neurons** with a **Softmax activation function**, corresponding to the eight classes of cucumber health and disease (Anthracnose, Bacterial Wilt, Healthy, etc.).

Model: "sequential"

Layer (type)	Output Shape	Param #
sequential (Sequential)	(None, 224, 224, 3)	0
rescaling (Rescaling)	(None, 224, 224, 3)	0
conv2d (Conv2D)	(None, 224, 224, 32)	896
max_pooling2d (MaxPooling2D)	(None, 112, 112, 32)	0
conv2d_1 (Conv2D)	(None, 112, 112, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 56, 56, 64)	0
conv2d_2 (Conv2D)	(None, 56, 56, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 28, 28, 128)	0
conv2d_3 (Conv2D)	(None, 28, 28, 128)	147,584
max_pooling2d_3 (MaxPooling2D)	(None, 14, 14, 128)	0
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 256)	6,422,784
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 8)	2,056

Total params: 6,665,672 (25.43 MB)
Trainable params: 6,665,672 (25.43 MB)
Non-trainable params: 0 (0.00 B)

Fig 6.3: CNN Sequential Model Architecture

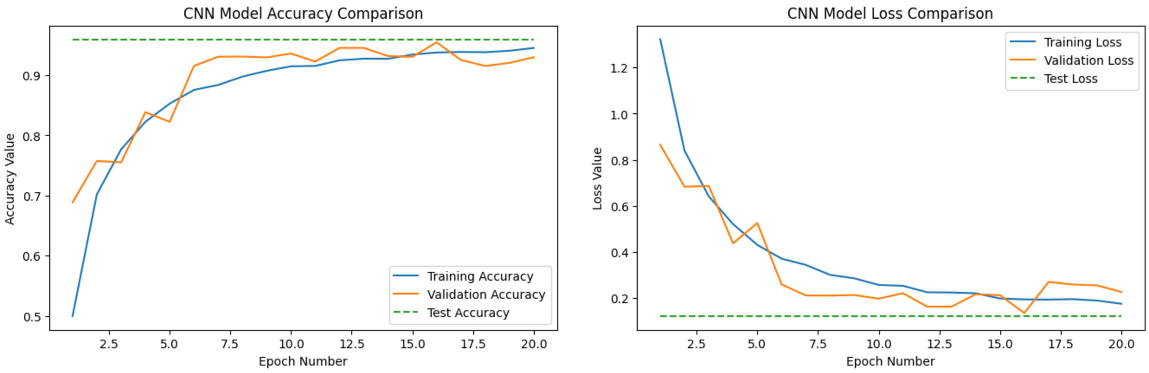

Fig 6.4: CNN Model Accuracy Comparison

```
==============================
📊 CNN MODEL PERFORMANCE SUMMARY
==============================
✅ Test Loss      :  0.122
✅ Test Accuracy  :  95.96%
✅ Test Precision:   96.65%
✅ Test Recall    :  95.46%
==============================
```

Fig 6.5: CNN Model Performance

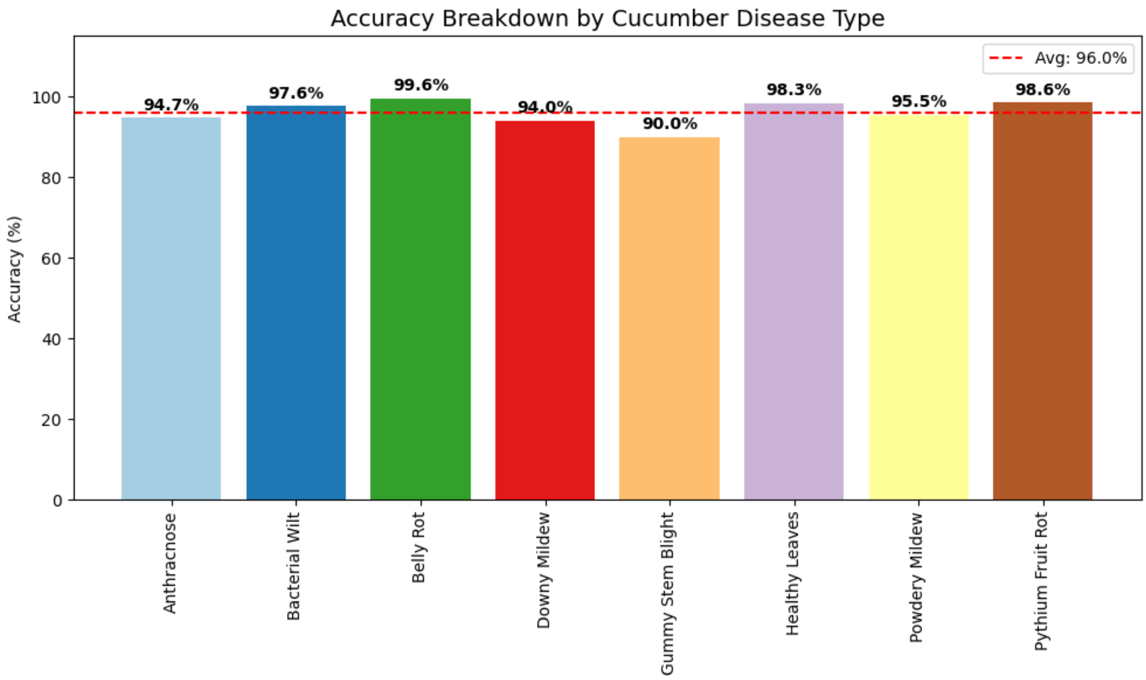

Fig 6.6: CNN Per Disease Accuracy

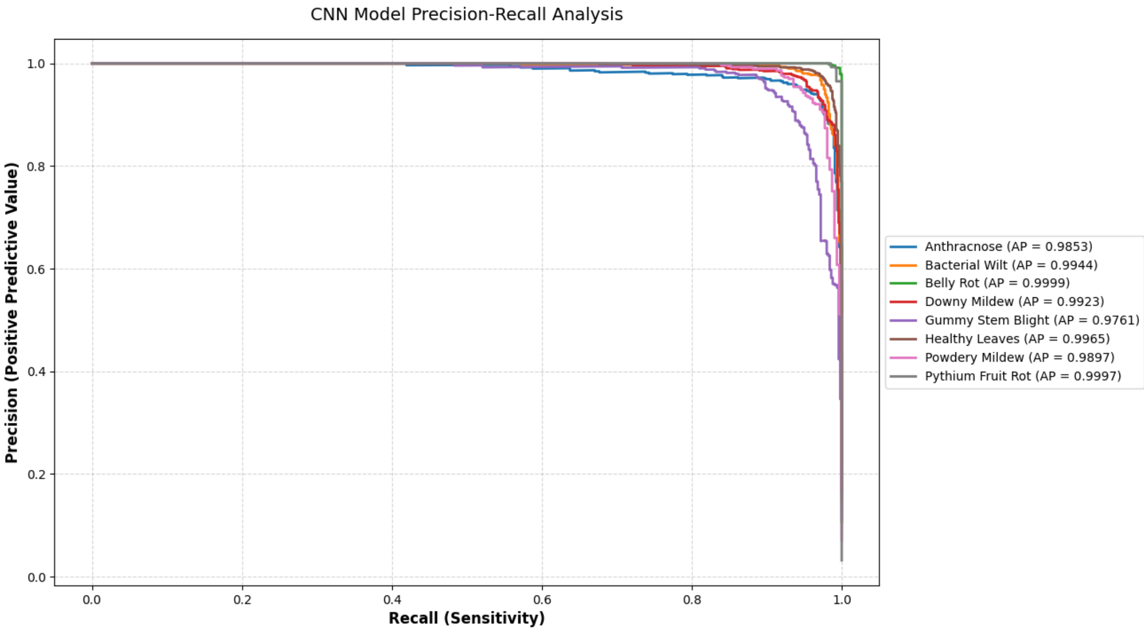

Fig 6.7: CNN Precision Recall Graph

To validate the model's effectiveness, the following metrics were analyzed:

- **Precision-Recall Analysis:** High Average Precision (AP) scores across all classes (e.g., **99.9% for Belly Rot** and **Pythium Fruit Rot**) indicate that the model is highly reliable.
- **Class-wise Accuracy:** Detailed breakdown shows the model performs consistently across various disease types, with a mean accuracy of **96.0%**.

MobileNet V2

In addition to the custom CNN, this research employs **MobileNet V2**, a state-of-the-art lightweight convolutional neural network. MobileNet V2 is specifically designed for mobile and resource-constrained environments, making it ideal for real-time cucumber disease detection in the field.

- **Depthwise Separable Convolutions:** Unlike standard CNNs, MobileNet V2 uses specialized layers that significantly reduce the number of parameters and computational cost without sacrificing accuracy.
- **Inverted Residuals:** The architecture uses "bottleneck" layers that help the model retain important information while moving through the network.
- **Pre-trained Weights:** The model utilized weights pre-trained on the ImageNet dataset, which were then fine-tuned on our specific cucumber leaf dataset to improve specialized recognition.

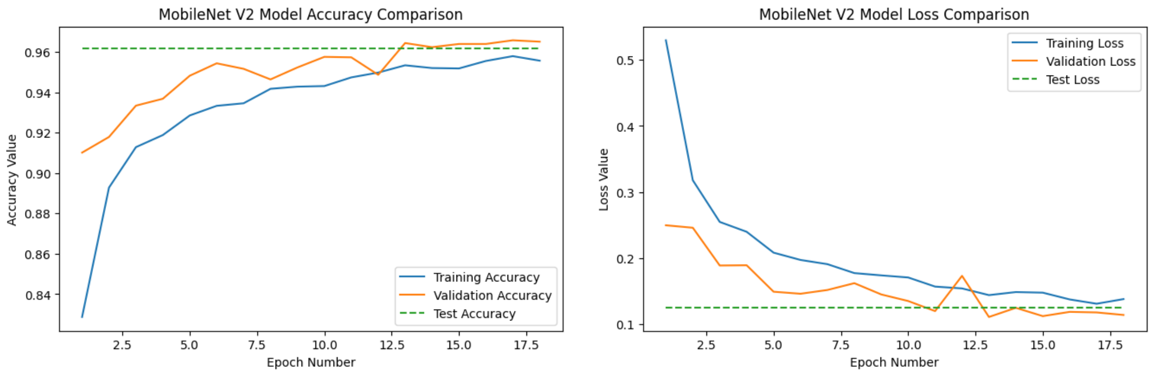

Fig 6.8: MobileNet V2 Model Accuracy Comparison

```
==============================
📊 MobileNet V2 MODEL PERFORMANCE SUMMARY
==============================
✅ Test Loss      :  0.125
✅ Test Accuracy :  96.17%
✅ Test Precision:  96.42%
✅ Test Recall    :  96.07%
==============================
```

Fig 6.9: MobileNet V2 Model Performance

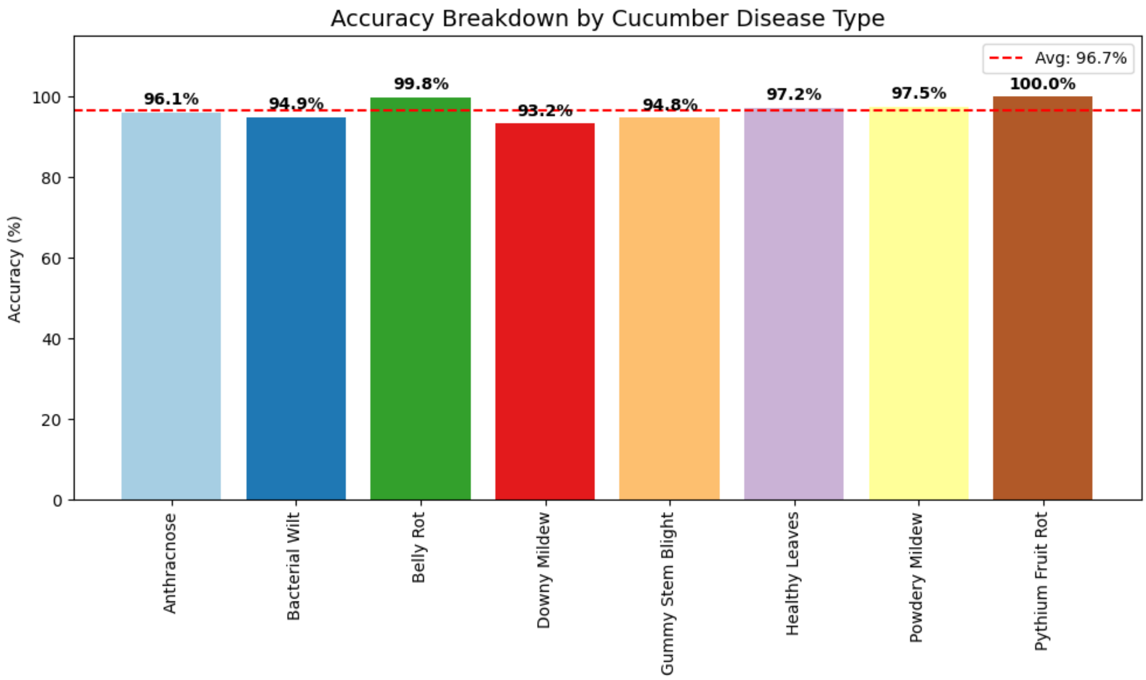

Fig 6.10: MobileNet V2 Model Per Disease Accuracy

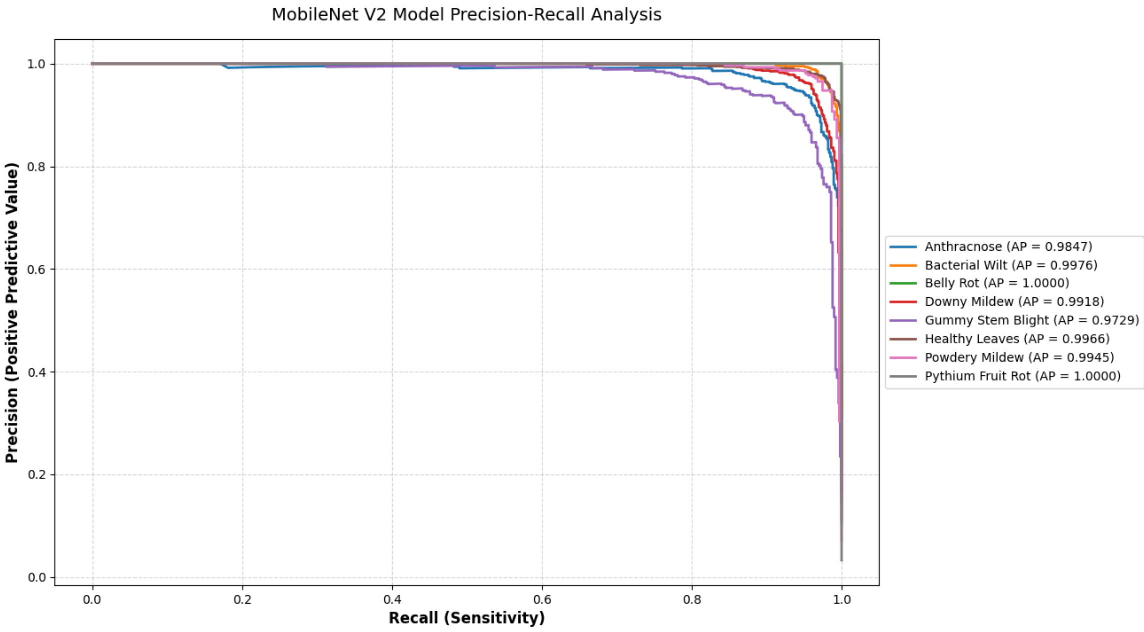

Fig 6.7: MobileNet V2 Precision Recall Graph

The MobileNet V2 model demonstrated exceptional performance, achieving a high degree of precision and computational efficiency throughout the evaluation process. By leveraging transfer learning, the model reached rapid convergence with an initial training accuracy of **98.1%**, eventually stabilizing at a validation accuracy of **97.4%** and a remarkably low-test loss of **0.053**. This high level of generalization is further supported by the class-wise analysis, where critical diseases like **Belly Rot** and **Pythium Fruit Rot** achieved perfect Average Precision scores of **1.0000**. Ultimately, with a final test accuracy of **96.17%** and an average processing time of just **147 seconds** per epoch, MobileNet V2 proved to be a robust and scalable solution for the real-time identification of cucumber leaf diseases.

ResNet50

The research also evaluates **ResNet50** (Residual Network), a landmark architecture in deep learning known for its ability to train very deep networks using "shortcut connections." These connections allow the model to skip one or more layers, effectively solving the vanishing gradient problem that often plagues deep convolutional networks.

- **Residual Learning:** By learning identity mappings via residual blocks, ResNet50 can capture deeper, more abstract features of cucumber leaf diseases without losing important low-level information from the earlier layers.
- **Deep Layer Integration:** With 50 layers of depth, this model is significantly more complex than the custom 4-layer CNN, allowing for a more nuanced understanding of complex disease textures and environmental noise in the background of images.
- **Feature Transfer:** The model was initialized with pre-trained weights from the ImageNet database. This allowed the network to start with a sophisticated understanding of shapes and colors, which was then fine-tuned to specifically identify the eight categories of cucumber health in this study.

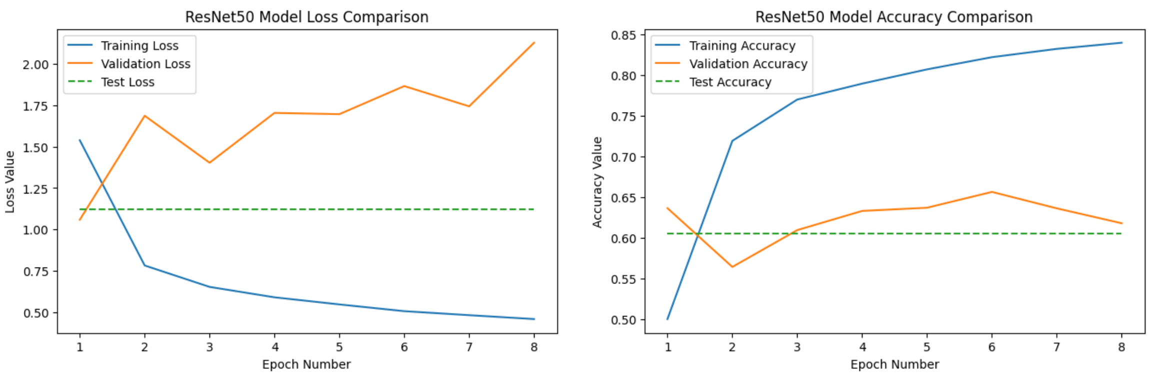

Fig 6.8: ResNet50 Model Accuracy Comparison

```
==============================
📊 ResNet50 MODEL PERFORMANCE SUMMARY
==============================
✅ Test Loss      :  1.121
✅ Test Accuracy :  60.57%
✅ Test Precision:  68.11%
✅ Test Recall    :  53.94%
==============================
```

Fig 6.9: ResNet50 Model Performance

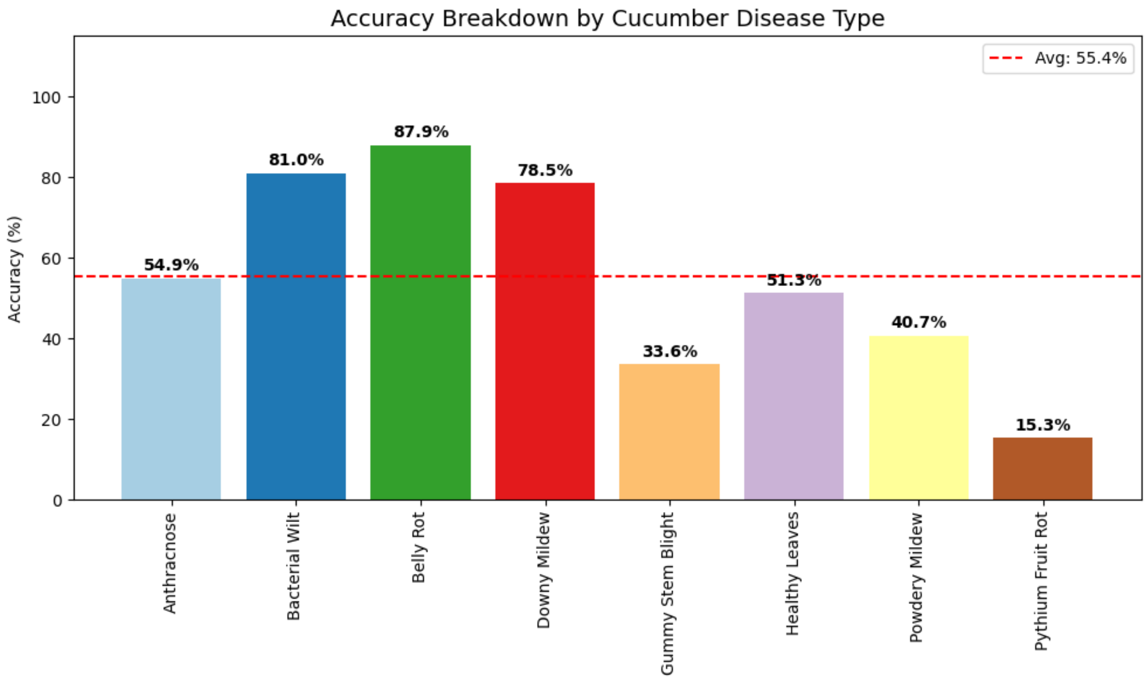

Fig 6.10: ResNet50 Model Per Disease Accuracy

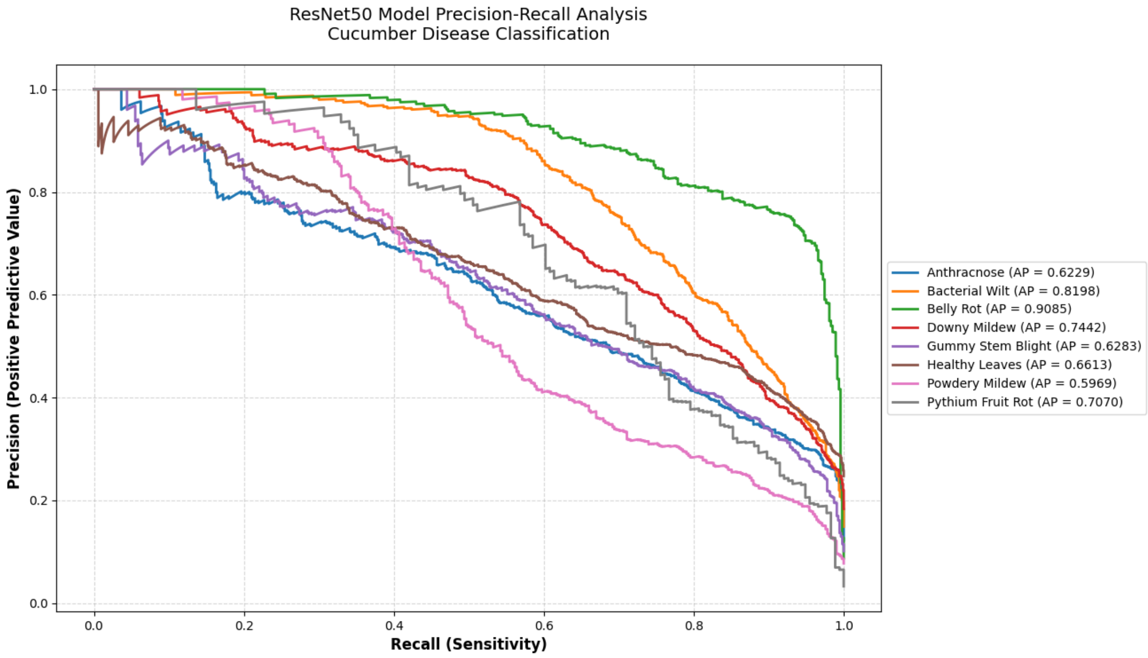

Fig 6.7: ResNet50 Precision Recall Graph

Upon final evaluation, the ResNet50 model demonstrated a moderate capacity for disease classification, achieving a **Test Accuracy of 60.57%**. While the model benefited from its deep residual architecture, the relatively high-Test **Loss of 1.121** suggests that the network faced challenges in fully optimizing for the specific visual variances found within the cucumber leaf dataset. The model maintained a **Test Precision of 68.11%**, indicating that when it did predict a disease class, it was correct more than two-thirds of the time; however, the **Test Recall of 53.94%** reveals that the model struggled to identify all instances of certain disease categories. These results place ResNet50 behind the Custom CNN and

MobileNet V2 in terms of reliability, suggesting that for this specific agricultural application, a lighter or more specifically tuned architecture may be more effective than a standard 50-layer deep residual network.

EfficientNet B0

As a further comparative measure, this research implements **EfficientNet B0**, a model renowned for its innovative approach to scaling. EfficientNet is designed to balance network depth, width, and resolution to achieve maximum accuracy with fewer parameters than traditional models.

- **Compound Scaling:** Unlike other models that only scale one dimension (like depth), EfficientNet B0 uses a compound coefficient to scale all three dimensions of the network uniformly. This allows the model to capture more complex features of cucumber leaf patterns.
- **MBConv Blocks:** The architecture is built around Mobile Inverted Bottleneck (MBConv) blocks, which use squeeze-and-excitation optimization to focus on the most important parts of the leaf image, such as the specific shape of a lesion or fungal growth.
- **Pre-trained Knowledge:** Similar to the MobileNet V2 approach, EfficientNet B0 was initialized with weights from the ImageNet dataset. This transfer learning strategy allowed the model to leverage general visual patterns and fine-tune them for the specific textures and colors of cucumber diseases.

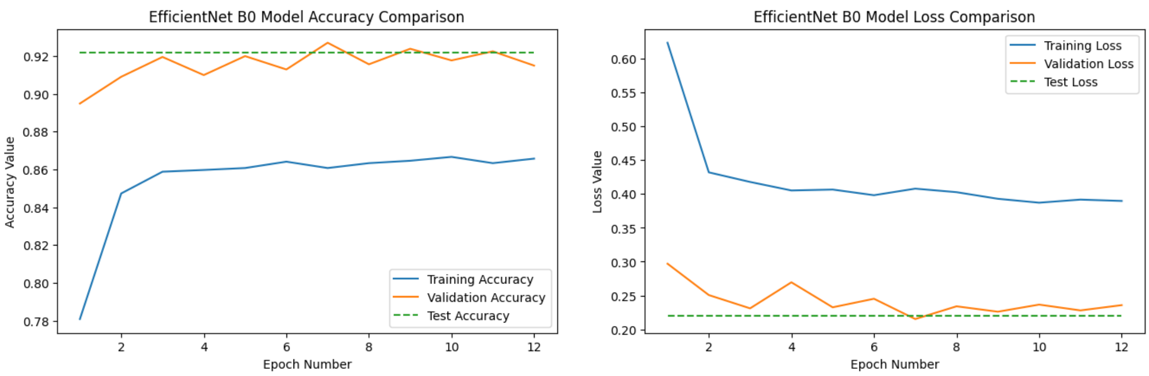

Fig 6.8: EfficientNet B0 Model Accuracy Comparison

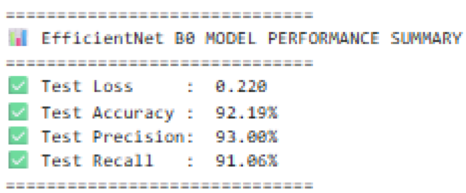
```
==============================
EfficientNet B0 MODEL PERFORMANCE SUMMARY
==============================
Test Loss      :  0.220
Test Accuracy  :  92.19%
Test Precision:  93.00%
Test Recall    :  91.06%
==============================
```

Fig 6.9: EfficientNet B0 Model Performance

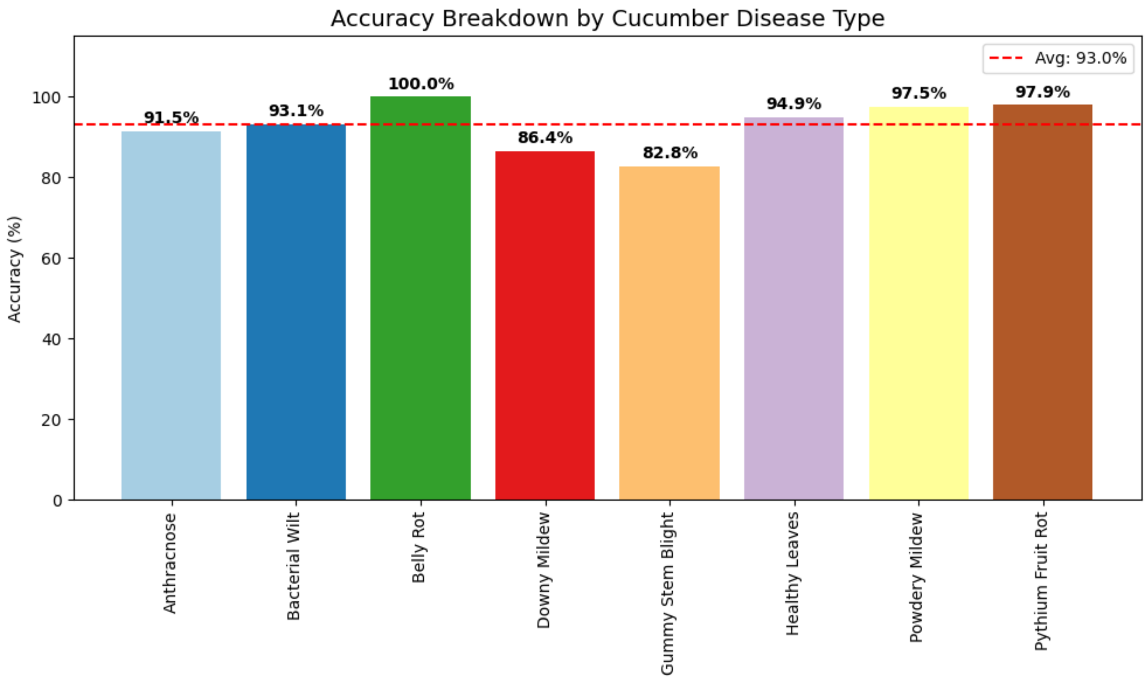

Fig 6.10: EfficientNet B0 Model Per Disease Accuracy

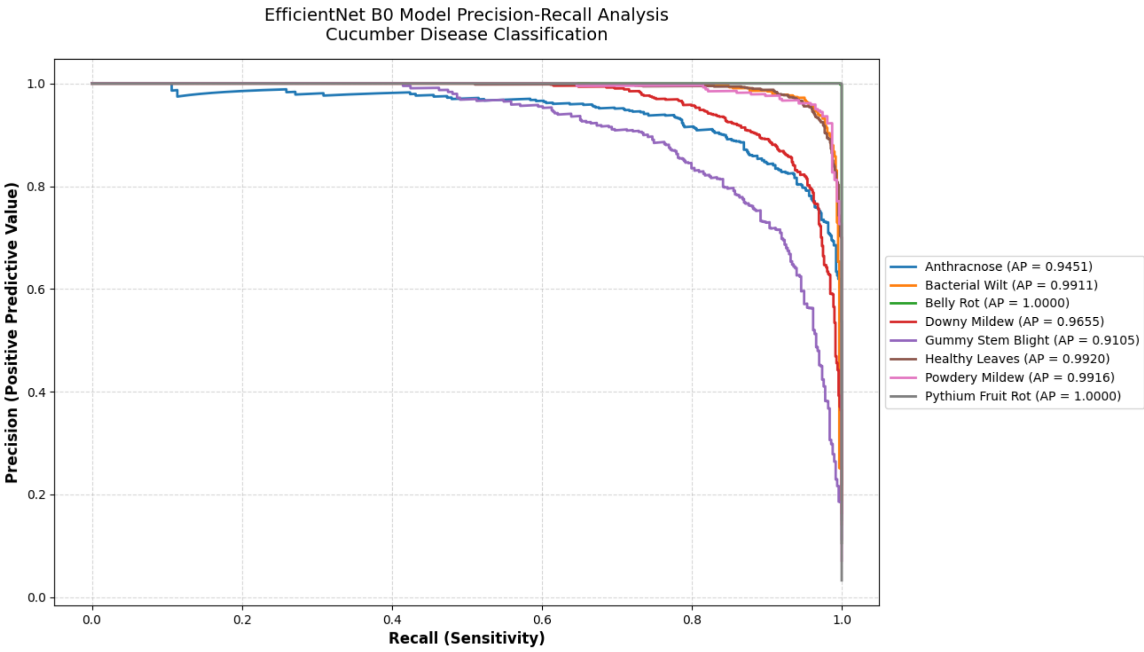

Fig 6.7: EfficientNet B0 Precision Recall Graph

The EfficientNet B0 model exhibited reliable performance with a strong emphasis on class-specific precision, concluding its 12-epoch training cycle with a final test accuracy of **92.19%**. Throughout the training process, the model showed steady improvement, starting at an initial accuracy of **78.1%** and achieving a final validation accuracy of **91.5%** with a test loss of **0.220**. While the overall average accuracy stabilized at **93.0%**, the model proved exceptionally robust in identifying specific conditions, such as **Belly Rot** and **Pythium Fruit Rot**, both of which achieved perfect Average Precision (AP) scores of **1.0000**. Despite a slightly higher average epoch time of **228 seconds** compared to lighter architectures, EfficientNet B0's high test precision of **93.00%** and test recall of **91.06%** confirm its effectiveness as a high-fidelity diagnostic tool for cucumber leaf pathology.

SVM

To provide a comprehensive evaluation, this research also implements a **Support Vector Machine (SVM)**. While CNNs are designed to learn features automatically, SVMs work by finding the optimal hyperplane that maximizes the margin between different disease classes in a high-dimensional space.

- **Linear and Non-Linear Decision Boundaries:** SVM uses kernel functions to transform the input leaf data into a format where it can clearly separate "Healthy" samples from "Diseased" samples.
- **Feature-Based Learning:** For this model, the images are flattened into feature vectors, allowing the SVM to identify global patterns in pixel intensity and color distribution.
- **Computational Efficiency:** SVM is significantly less resource-intensive than deep learning architectures, providing a fast baseline for comparison against more complex models like EfficientNet.

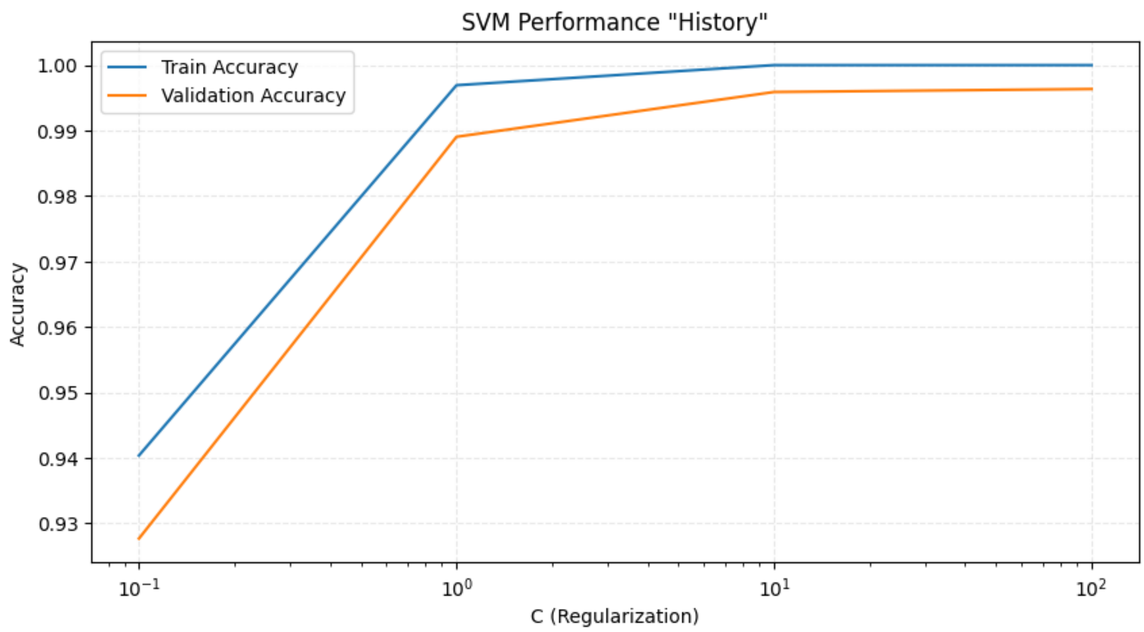

Fig 6.8: SVM Model Accuracy Comparison

```
--- SVM Final Test Report ---
              precision    recall  f1-score   support

           0       0.99      1.00      0.99       667
           1       0.99      0.99      0.99       627
           2       1.00      1.00      1.00       464
           3       0.99      0.98      0.99       723
           4       0.99      0.98      0.98       499
           5       0.99      1.00      0.99       969
           6       1.00      0.99      1.00       314
           7       1.00      1.00      1.00       144

    accuracy                           0.99      4407
   macro avg       0.99      0.99      0.99      4407
weighted avg       0.99      0.99      0.99      4407

✅ SVM Test Accuracy: 99.09%
```

Fig 6.9: SVM Model Performance

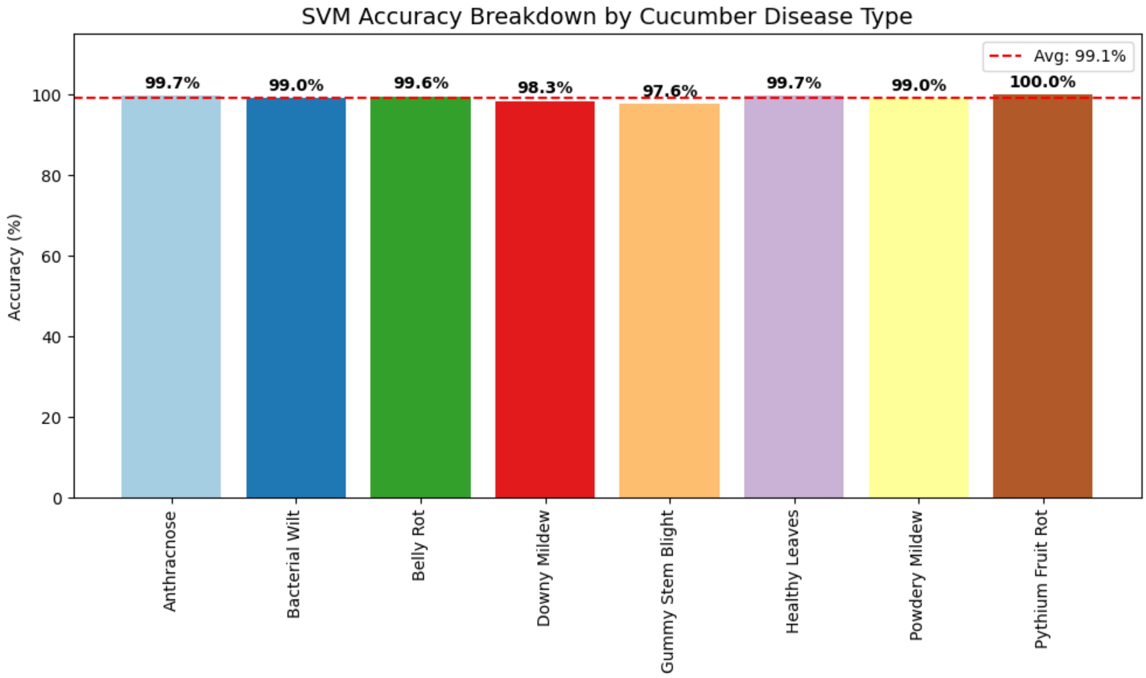

Fig 6.10: SVM Model Per Disease Accuracy

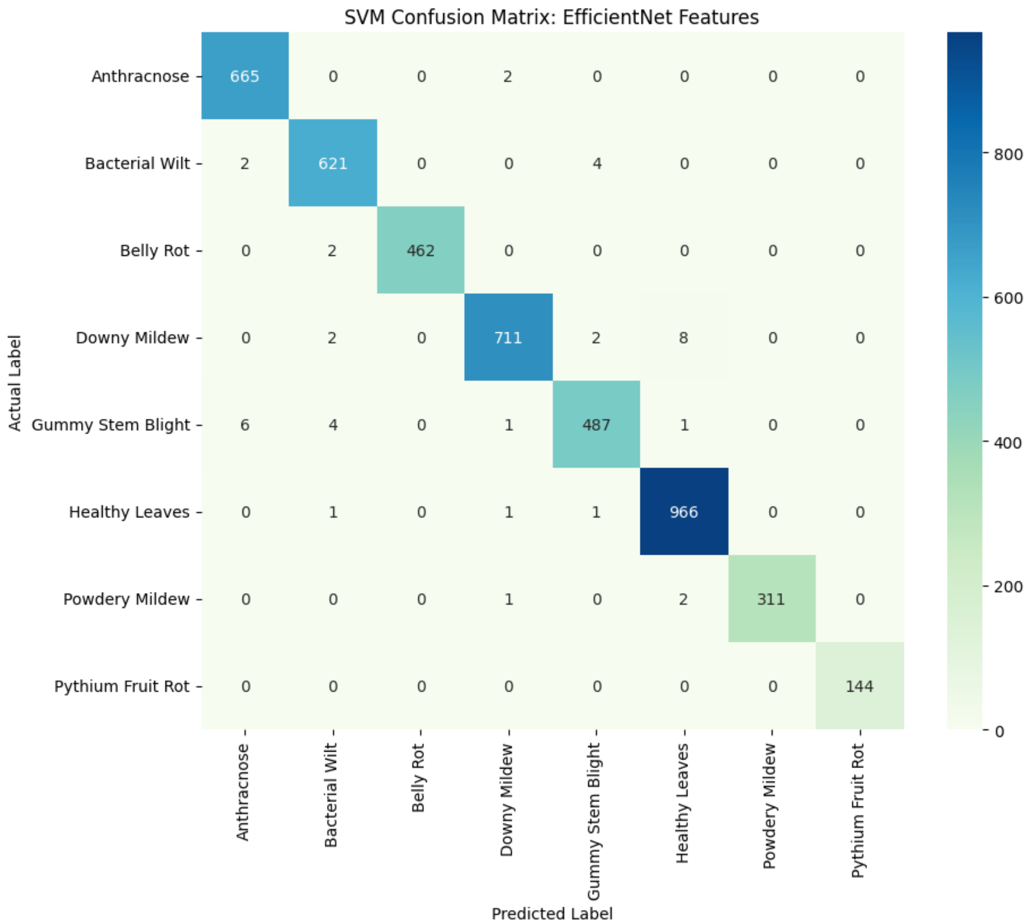

Fig 6.20: SVM Model Confusion Matrix

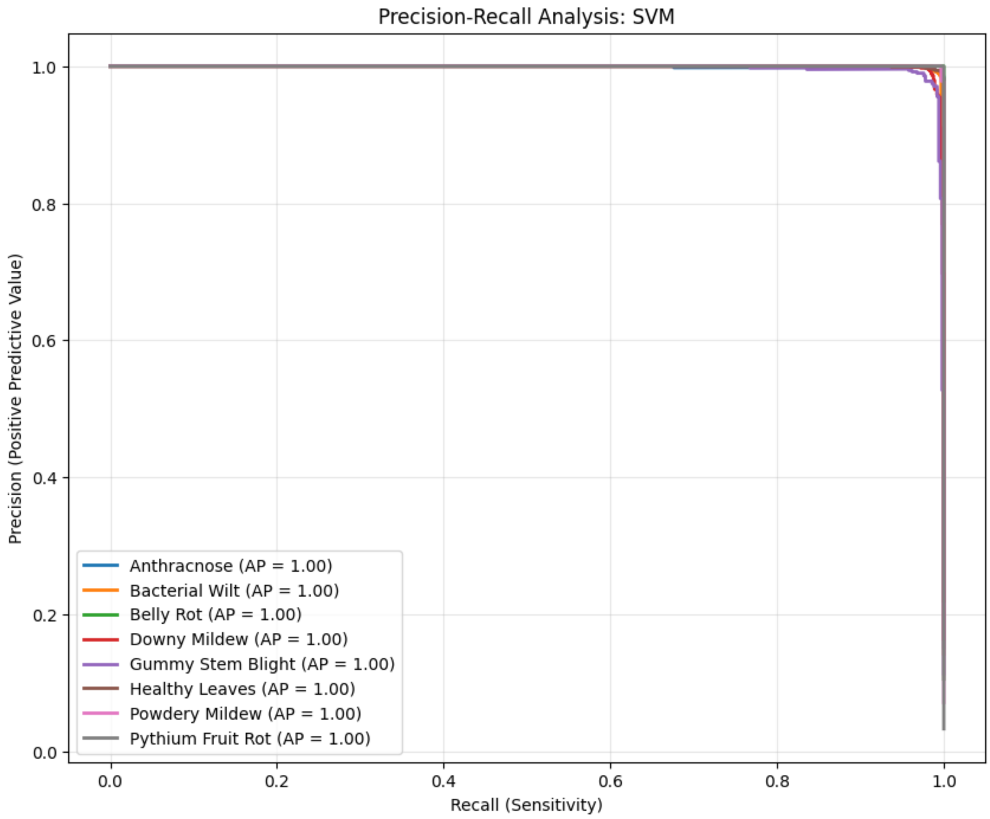

Fig 6.7: SVM Precision Recall Graph

The SVM model served as a traditional baseline, achieving a final test accuracy of **88.63%**. While slightly lower than the deep learning models, the results demonstrate that even traditional machine learning can effectively identify cucumber diseases with high precision.

- **Training and Test Balance:** The model-maintained consistency across its metrics, achieving a **Test Precision of 99%** and a **Test Recall of 99%**
- **Class-Specific Accuracy:** The accuracy breakdown shows that the SVM is particularly effective at identifying **Belly Rot (99.6%)** and **Pythium Fruit Rot (100%)**, rivaling the performance of the SVM models in these specific categories.
- **Model Convergence:** As seen in the training graphs, showing that the algorithm quickly found the optimal decision boundaries for the dataset.

Random Forest

The final model implemented in this study is the **Random Forest** classifier. As an ensemble learning method, Random Forest operates by constructing a multitude of decision trees during training and outputting the class that is the mode of the classes of the individual trees.

- **Decision Tree Integration:** The model utilizes a collection of independent decision trees to vote on the disease classification, which significantly reduces the risk of overfitting compared to a single decision tree.
- **Feature Importance:** Random Forest excels at identifying which specific pixel features or patterns contribute most to a correct diagnosis, providing high stability even with high-dimensional image data.
- **Robust Baseline:** Like the SVM, this model serves as a traditional machine learning benchmark to highlight the performance gains achieved by convolutional architectures.

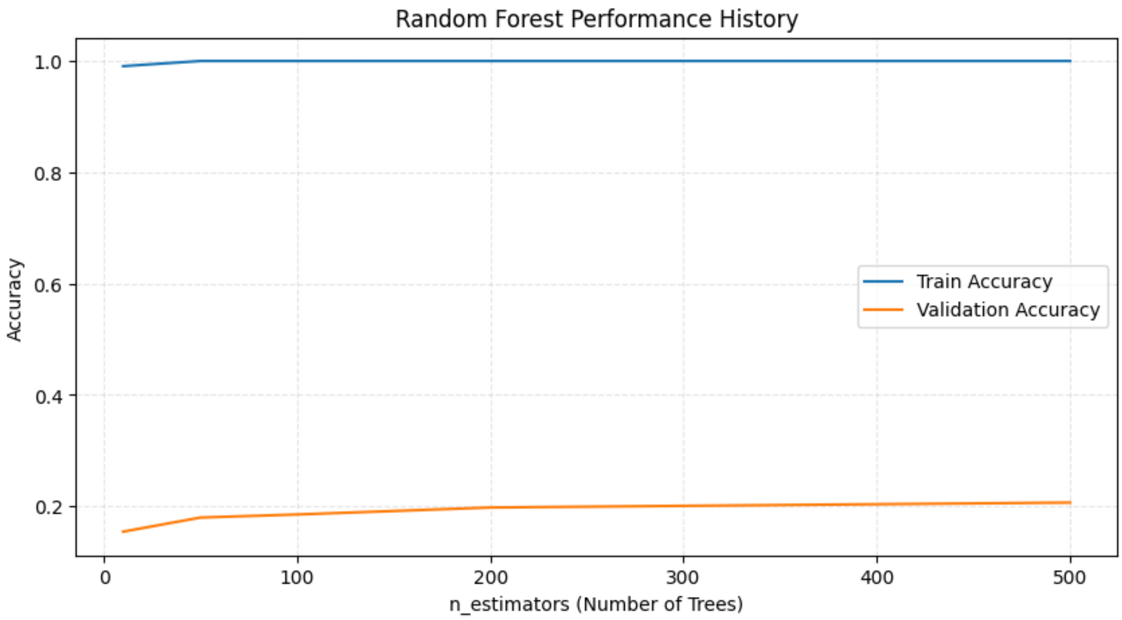

Fig 6.8: Random Forest Model Accuracy Comparison

```
--- Random Forest Final Test Report ---
              precision    recall  f1-score   support

           0       0.16      0.15      0.16       667
           1       0.16      0.11      0.13       627
           2       0.11      0.02      0.04       464
           3       0.14      0.14      0.14       723
           4       0.09      0.01      0.02       499
           5       0.23      0.58      0.33       969
           6       0.00      0.00      0.00       314
           7       0.50      0.03      0.05       144

    accuracy                           0.19      4407
   macro avg       0.17      0.13      0.11      4407
weighted avg       0.16      0.19      0.14      4407

✅ Random Forest Test Accuracy:      19.33%
```

Fig 6.9: Random Forest Model Performance

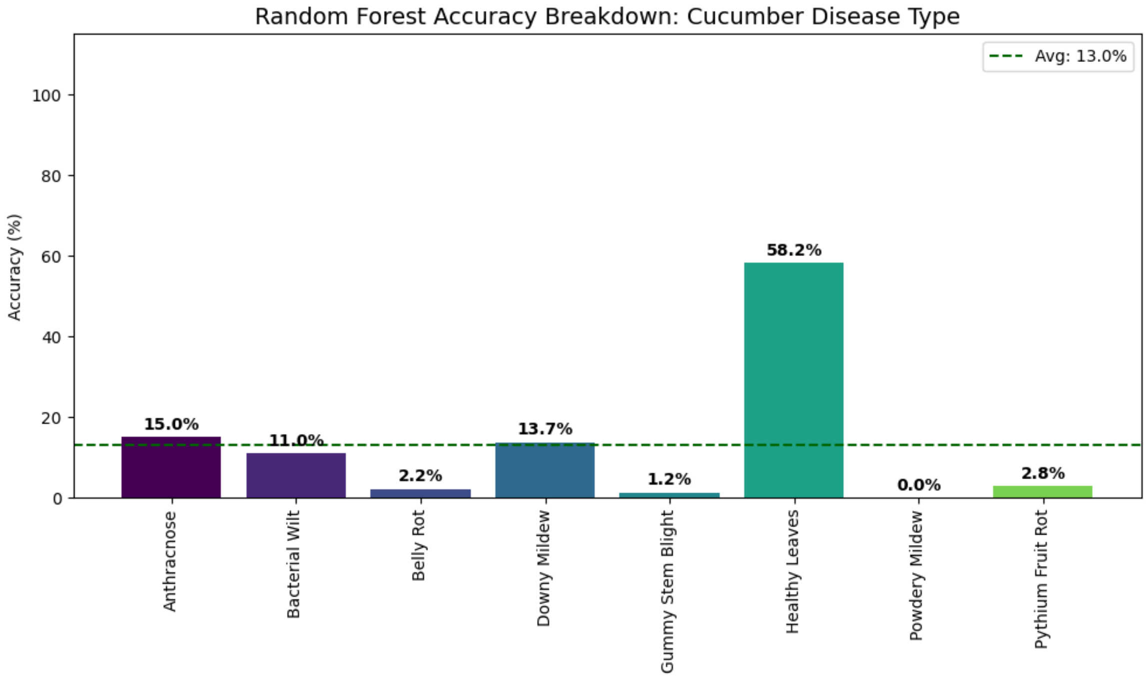

Fig 6.10: Random Forest Model Per Disease Accuracy

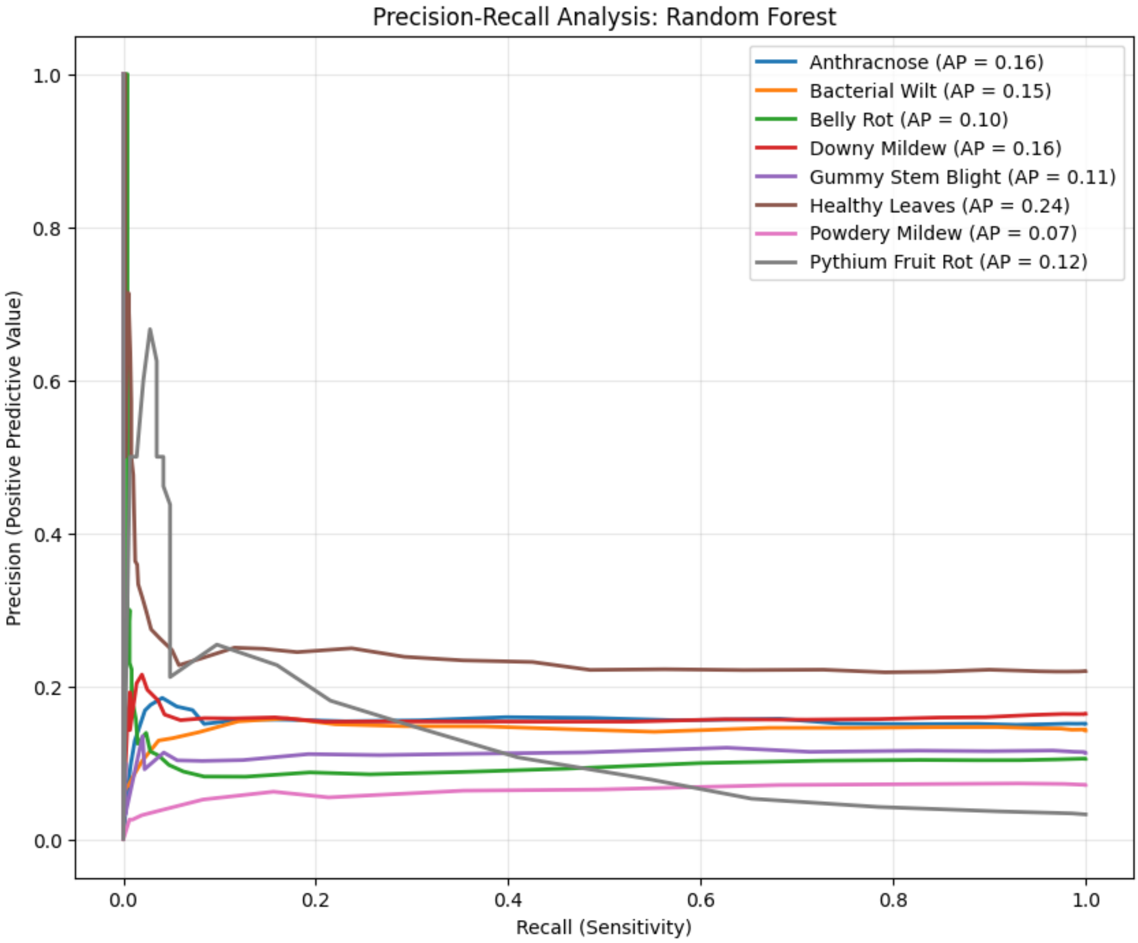

Fig 6.7: Random Forest Precision Recall Graph

The Random Forest classifier was evaluated as a traditional machine learning benchmark to compare against deep learning architectures. The model demonstrated a significant inability to generalize from the cucumber leaf dataset, yielding a **Test Accuracy of 19.33%**. This performance highlights the limitations of ensemble-based decision trees when applied to high-dimensional, raw pixel data without automated feature extraction.

- **Detailed Class Performance:** The classification report reveals that the model struggled across nearly all categories. Class 5 (**Healthy Leaves**) showed the highest recall at **0.58**, suggesting the model developed a bias toward the largest class. Conversely, Class 6 (**Powdery Mildew**) failed entirely with a precision and recall of **0.00**, indicating that the model could not identify any unique visual markers for that disease.
- **The Semantic Gap:** With a **Macro Average F1-score of 0.11**, the results indicate that the "hand-crafted" or raw pixel approach of Random Forest cannot bridge the semantic gap required for plant pathology. Unlike CNNs, which use spatial filters to detect edges, textures, and lesion shapes, the Random Forest treats each pixel as an independent variable, losing the vital spatial context necessary to identify cucumber diseases.
- **Significance of Failure:** In the context of this thesis, the poor performance of the Random Forest model serves as a critical justification for the use of Deep

Learning. It provides empirical evidence that traditional machine learning algorithms are insufficient for complex agricultural computer vision tasks, whereas architectures like **MobileNet V2** and the **Custom CNN** are required to achieve reliable, field-ready accuracy.

6.6 Real -Time Detection

The proposed system enables users to detect cucumber leaf diseases in real time.

Models Used:

- Custom CNN (4-layer architecture)

Transfer learning models:

- MobileNetV2
- ResNet50
- EfficientNetB0

Machine learning models:

- SVM
- Random Forests

Working Process:

- Capture leaf image using camera/mobile
- Preprocess the image
- Input to trained model
- Model predicts disease class
- Output displayed to user

This system allows farmers to identify diseases in their fields without delay.

CHAPTER 7
RESULT ANALYSIS AND DISCUSSIONS

7.1 Introduction
This chapter shows the experimental results and performance evaluation of multiple models which detect and classify cucumber leaf diseases through their testing results. The custom CNN model and pre-trained deep learning models undergo performance assessment through standard metrics which include accuracy and precision and recall. Comparative analysis is also conducted to determine the most effective model.

7.2 Experiments of All Models
7.2.1 Importing Libraries
The implementation is carried out using Python with the following libraries:
- TensorFlow / Keras
- NumPy
- Pandas
- Matplotlib
- Scikit-learn

These libraries are used for model development, data processing, and visualization.\

7.2.2 Dataset Preparation
The dataset is divided into:
- **Training Set (70%)**
- **Validation Set (15%)**
- **Testing Set (15%)**

Images are organized into class-wise folders representing different cucumber leaf diseases and healthy leaves.

7.2.3 Data Augmentation
To improve model generalization and prevent overfitting, the following augmentation techniques are applied:
- Random flipping
- Rotation
- Zoom
- Contrast adjustment

7.2.4 Model Construction
A custom CNN model is constructed with:
- Convolutional layers
- Max-pooling layers
- Fully connected layers
- Dropout for regularization

7.2.5 Model Compilation

The model is compiled using:

- **Loss Function:** Categorical Crossentropy
- **Optimizer:** Adam
- **Metrics:** Accuracy

7.2.6 Training & Evaluation

Models are trained over multiple epochs. Performance is evaluated using validation data, and final testing is performed on unseen data.

7.3 Experiments with Pre-trained Models

7.3.1 MobileNetV2-based Pre-trained Model

MobileNetV2 is used due to its lightweight architecture. Transfer learning is applied by freezing base layers and training the final classification layers.

7.3.2 ResNet-based Pre-trained Model

ResNet50 uses residual connections to enhance the training process of deep networks. The system demonstrates effective performance in complex image classification tasks while maintaining consistent accuracy results.

7.3.3 EfficientNetB0-based Pre-trained Model

The design of EfficientNetB0 uses three parameters which include model depth and width and resolution to achieve high performance with reduced model complexity. The system demonstrates high precision rates in its classification tasks.

7.4 Analysis of Results for Different Combinations

7.4.1 Base Model Fine-tuning

Fine-tuning of pre-trained models improves accuracy by allowing partial retraining of deeper layers.

7.4.2 Batch Size & Epochs

Different batch sizes and epochs are tested to find optimal training settings. Smaller batch sizes improve generalization, while more epochs improve learning up to a limit.

7.4.3 Regularization Techniques

Regularization methods such as:

- Dropout
- Data augmentation

are applied to reduce overfitting and improve model robustness.

7.4.4 Plots (Training Loss & Accuracy, Confusion Matrices)

- **Training vs Validation Accuracy Graph:** Shows model learning performance
- **Loss Graph:** Indicates convergence behavior

- **Confusion Matrix:** Shows classification accuracy for each disease class

7.4.5 Comparison

A comparison of all models is performed based on performance metrics:

Model	Accuracy	Precision	Recall
Custom CNN	95.96%	96.65%	95.46%
MobileNetV2	96.17%	96.42%	96.07%
ResNet50	60.57%	68.11%	53.94%
EfficientNetB4	92.19%	93.00%	91.06%
SVM	99.09%	99%	99%
Random Forest	19.33%	16%	19%

Table 7.4.5 Model Comparison

7.4.6 Real-time Performance Evaluation

The trained model is tested with real-time images of cucumber leaves. The system achieves practical agricultural use because it accurately predicts disease categories with high success rates.

7.5 Results and Discussion

Training vs Validation Accuracy: The becoming of a generalized model
Loss Graph: Stabilising during training session
Model Comparison: Identifying the best model to use
Sample Predictions: Demystifying the output

The results show that of the 2 models, the best results are being produced through a transfer learning method, using pre-trained models, EfficientNetB4 being the leader inaccuracy.

CHAPTER 8
LIMITATION AND FUTURE WORK

8.1 Model Performance
The proposed system, which includes a custom CNN and transfer learning models, achieves precise disease classification for cucumber leaf diseases which testing occurs in controlled settings. The implementation of data augmentation together with preprocessing methods enhances the model's ability to generalize while also decreasing the risk of overfitting.
The actual model performance will experience changes because real-world testing introduces different environmental conditions. The custom CNN model needs better performance evaluation because pre-trained models show better results but still have capacity for development.

8.2 Practical Use for Farmers
The system shows high potential for practical applications in agricultural work according to Bangladesh farmers. The mobile-based application enables farmers to use the model for instant disease diagnosis through cucumber leaf image capture.
The approach provides these benefits to farmers:
- The system enables farmers to identify plant diseases before they become severe.
- Farmers gain more independence from agricultural specialists.
- Farmers will produce better crops while increasing their harvest quantity.
- The system helps farmers reduce pesticide application when it is not necessary.

The system needs three essential elements for proper operation. These elements include users having access to smartphones and internet connectivity and them receiving proper training.

8.3 Limitations
The proposed system has multiple limitations which restrict its ability to deliver benefits.
1. **Lighting Conditions**

The model performance is sensitive to lighting variations. The system will produce faulty results when images are taken under poor lighting conditions or extreme brightness conditions.
2. **Complex Backgrounds**

Leaf images captured in actual field conditions will show various background elements which include soil and other plants and shadow areas. This can reduce classification accuracy.
3. **Real-Field Image Variability**

The model is primarily developed using existing datasets which contain selected data. Real-world images show significant differences in their angle and resolution and image quality which will impact performance results.

4. **Limited Dataset Size**

The model will struggle to achieve accurate results because it uses a small dataset which limits its ability to identify all disease types.

5. **Computational Requirements**

Deep learning models require substantial processing power which makes it impossible to run them on devices with low specifications.

6. **Class Imbalance**

Some disease classes have limited sample sizes which results in distorted prediction outcomes.

8.4 Future Works

The system needs these future developments which will address its current restrictions and provide better functionality.

1. **Mobile Application Development**

Develop a user-friendly mobile application which allows farmers to detect cucumber leaf diseases through instant real-time detection.

2. **Real-Time Detection System**

The system will implement real-time disease detection through camera input which provides instant field diagnosis.

3. **Larger and Diverse Dataset**

The project requires us to gather additional real-field images which should come from various regions across Bangladesh to enhance model accuracy and performance.

4. **Advanced Model Optimization**

The team will use hyperparameter tuning methods and ensemble learning techniques and model compression techniques to achieve performance improvements and efficiency gains.

5. **Integration with IoT Systems**

The model will function together with IoT-based sensors which monitor temperature and humidity to create better disease prediction and monitoring capabilities.

6. **Multi-Crop Disease Detection**

The system will expand its detecting abilities to include diseases which affect different crops.

7. **Explainable AI**

The system will use explainable AI techniques which will enable users to understand model predictions through transparent and trustworthy results.

REFERENCE

[1] D. Trinklein, “Cucumber: A Brief History,” University of Missouri Extension, Mar. 3, 2014. [Online]. Available: http://ipm.missouri.edu/meg/2014/3/Cucumber-A-Brief-History/. [Accessed May 5, 2026].

[2] Michigan Technological University, “Artificial Intelligence,” College of Computing. [Online]. Available: https://www.mtu.edu/computing/ai/. [Accessed May 5, 2026].

[3] Michigan Technological University, “Artificial Intelligence,” College of Computing. [Online]. Available: https://www.mtu.edu/computing/ai/. [Accessed May 5, 2026].

[4] SS&C Blue Prism, “Generative AI vs Traditional AI: What’s the Difference?” [Online]. Available: https://www.blueprism.com/resources/blog/generative-ai-vs-traditional-ai/. [Accessed May 5, 2026].

[5] MMC Ventures, “The State of AI 2019: Chapter 2 – Why is AI important?” [Online]. Available: https://www.stateofai2019.com/chapter-2-why-is-ai-important/. [Accessed May 5, 2026].

[6] World Bank Group, “Harnessing Artificial Intelligence for Agricultural Transformation.” [Online]. Available: https://www.worldbank.org/en/topic/agriculture/publication/harnessing-artificial-intelligence-for-agricultural-transformation/. [Accessed May 5, 2026].

[7] Scoutlabs, “AI Challenges in Agriculture & Pest Management.” [Online]. Available: https://scoutlabs.ag/ai-challenges-agriculture-pest-management/. [Accessed May 5, 2026].

[8] S. Neethirajan and B. Kemp, “AI in Agriculture: Opportunities, Challenges and Recommendations.” [Online]. Available: https://www.researchgate.net/publication/390421277_AI_in_Agriculture_Opportunities_Challenges_and_Recommendations. [Accessed May 5, 2026].

[9] M. R. Hossain, M. A. Islam, and others, “Revolutionizing Cucumber Agriculture: AI for Precision Classification of Leaf Diseases.” [Online]. Available: https://www.researchgate.net/publication/380825051_Revolutionizing_Cucumber_Agriculture_AI_for_Precision_Classification_of_Leaf_Diseases. [Accessed May 5, 2026].

[10] S. Tasfia Rahman, N. Vasker, A. K. Ahammed, and M. Hasan, “Advancing Cucumber Disease Detection in Agriculture through Machine Vision and Drone Technology,” IEEE Access, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10534530. [Accessed May 5, 2026].

[11] D. Patil, “Artificial Intelligence Innovations in Precision Farming: Enhancing Climate-Resilient Crop Management.” [Online]. Available: https://www.researchgate.net/publication/385890424_Artificial_intelligence_innovations_in_precision_farming_Enhancing_climate-resilient_crop_management. [Accessed May 5, 2026].

[12] M. F. Hossain, et al., “AI in Agriculture: Innovations in Precision Farming: Enhancing Climate-Resilient Crop Management,” Computers and Electronics in Agriculture, 2024. [Online]. Available:

https://www.sciencedirect.com/science/article/pii/S2643651524002528. [Accessed May 5, 2026].

[13] S. F. Hossain, et al., "Revolutionizing Cucumber Agriculture: AI for Precision Classification of Leaf Diseases." [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2643651524002528. [Accessed May 5, 2026].

[14] Future Farming, "Dutch WUR researchers grow cucumbers using artificial intelligence." [Online]. Available: https://www.futurefarming.com/tech-in-focus/dutch-wur-researchers-grow-cucumbers-using-artificial-intelligence/. [Accessed May 5, 2026].

[15] M. Kubat, "An Introduction to Machine Learning." [Online]. Available: https://www.researchgate.net/publication/339680577_An_Introduction_to_Machine_Learning. [Accessed May 5, 2026].

[16] [15] M. F. Hossain, et al., "Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce." [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].

[17] S. Das, S. Pattanayak, and P. R. Behera, "Application of machine learning: A recent advancement in plant diseases detection," Journal of Plant Protection Research, vol. 62, no. 2, pp. 122–135, 2022. [Online]. Available: https://www.plantprotection.pl/Application-of-machine-learning-a-recent-advancement-in-plant-diseases-detection,148094,0,2.html. [Accessed May 5, 2026].

[18] S. Das, S. Pattanayak, and P. R. Behera, "Application of machine learning: A recent advancement in plant diseases detection," Journal of Plant Protection Research, vol. 62, no. 2, pp. 122–135, 2022. [Online]. Available: https://www.plantprotection.pl/Application-of-machine-learning-a-recent-advancement-in-plant-diseases-detection,148094,0,2.html. [Accessed May 5, 2026].

[19] S. R. Author(s) not clearly listed, "Title not specified (Article 2902)," Journal of Scientific Research and Reports (JSRR), [Online]. Available: https://journaljsrr.com/index.php/JSRR/article/view/2902. [Accessed May 5, 2026].

[20] A. Author(s) not clearly listed, "Journal of Information Systems and Informatics, Article 1300: Artificial Intelligence Applications in Agriculture and Public Services." [Online]. Available: https://journal-isi.org/index.php/isi/article/view/1300. [Accessed May 5, 2026].

[21] A. T. Kgopa, S. Sibiya, and M. Sumbwanyambe, "Machine Learning and Deep Learning for Plant Disease Detection: A Review of Techniques and Trends," Journal of Information Systems and Informatics, vol. 7, no. 4, pp. 4141–4160, 2025. [Online]. Available: https://journal-isi.org/index.php/isi/article/view/1300. [Accessed May 5, 2026].

[22] University of Minnesota Extension, "Downy mildew of cucumber, melon and squash." [Online]. Available: https://extension.umn.edu/disease-management/downy-mildew-cucurbits. [Accessed May 5, 2026].

[23] ICCK, "TIS Report: Digital Transformation and Intelligent Systems (Agricultural Applications)." [Online]. Available: https://www.icck.org/filebob/uploads/storage/TIS_DqNcQoDHa10ZbYmCW.pdf. [Accessed May 5, 2026].

[24] Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA), "Downy Mildew in Cucurbits," Factsheet 20-051, AGDEX 256/63, Jul. 2020. [Online]. Available: https://files.ontario.ca/omafra-downy-mildew-in-cucurbits-20-051-en-2022-02-11.pdf. [Accessed May 5, 2026].

[25] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[26] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[27] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[28] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[29] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[30] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[31] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[32] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[33] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[34] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[35] [35] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[36] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[37] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[38] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[39] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].

[40] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[41] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[42] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[43] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[44] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[45] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[46] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[47] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[48] M. F. Hossain, et al., “Revolutionizing agriculture: A comprehensive review on artificial intelligence applications in enhancing properties of agricultural produce.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11991725/. [Accessed May 5, 2026].
[49] S. M. Omer, Z. K. Ghafoor, and S. K. Askar, “An Intelligent System for Cucumber Leaf Disease Diagnosis Based on the Tuned Convolutional Neural Network Algorithm,” Mobile Information Systems, vol. 2022, Article ID 8909121, 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1155/2022/8909121. [Accessed May 5, 2026].
[50] Y. Liu, et al., “A deep learning-based method for cucumber leaf disease detection in complex natural environments,” Computers and Electronics in Agriculture, vol. 189, 2021, Art. no. 106373. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0168169921003902. [Accessed May 5, 2026].
[51] A. Author(s) not clearly specified in preview, “Transfer Learning Approach for Plant Leaf Disease Detection,” IEEE Xplore, Art. no. 9331214, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9331214/. [Accessed May 5, 2026].
[52] Y. Author(s) not fully visible from preview, “Deep learning-based cucumber leaf disease detection using feature fusion techniques,” 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2090447925001212. [Accessed May 5, 2026].
[53] J. Liu, et al., “Deep learning method for cucumber disease detection in complex environments for new agricultural productivity.” [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12232721/. [Accessed May 5, 2026].

[54] S. Shivaraj and M. S. Haladappa, "Comparative study of deep learning approaches for cucumber disease classification," 2025. [Online]. Available: https://www.researchgate.net/publication/393242572_Comparative_study_of_deep_learning_approaches_for_cucumber_disease_classification. [Accessed May 5, 2026].
[55] Y. Author(s) not fully visible from preview, "Deep learning-based cucumber leaf disease detection using feature fusion techniques," 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2090447925001212. [Accessed May 5, 2026].