Plant Disease Detection System

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning
with
TechSaksham – A joint CSR initiative of Microsoft & SAP

by

Karan , kjay02689@gmail.com

Under the Guidance of

P.Raja , Master Trainer , Edunet Foundation

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deep sense of thankfulness to each and every co-ordinator of this AICTE-Internship who helped us either directly or indirectly.

I would like to thank the Edunet Foundation for their invaluable mentorship and guidance over the past few weeks. I also Extend my Gratitude to Mr.P.Raja for his wonderful mentoring in this project. His unwavering support has been a big part in my professional and personal growth.

One of the significant lesson I have Learned Through These sessions is that I can Enhance my skills in Artificial Intelligence and Machine Learning Combining Together with the Agriculture field and specifically in the detection of diseases in the plants.

ABSTRACT

Plant disease and pest detection is a very important research content in the field of machine vision. It is a technology that uses machine vision equipments to acquire images to judge whether they are disease and pest in the collected plant images.

As we all know that the Plant which is Healthier can be easily affected by the any of the insects or diseases that are spread from any other bacterium and virus.This may lead to Economical Disruptions , Reducing Food Quality and Lose Of Crops.

The Main Objective of this Project is to create a detection system that detects the disease in the plants through scanning the leaves and reducing the crop loses. The other objective is to provide farmers a user friendly tool to detect disease and reducing the efforts required for manual inspection.

The methodology for making such detection system comes under the various process like the following

- Data Collection
- Image Preprocessing
- Feature Extraction
- Model Training
- Model Evaluation
- Deployment

The Main Key Result are Like the Data Has the High Accuracy , Early Detection of diseases , Improved Crop Yields and Scalablity.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Problem Statement:

As we all know that the Plant which is Healthier can be easily affected by the any of the insects or diseases that are spread from any other bacterium and virus.This may lead to Economical Disruptions , Reducing Food Quality and Lose Of Crops and even to the food scarcity.

To create a trustworthy and accurate system that can automatically detect and diagnose plant illnesses in a quick and timely manner is the other main problem statement for plant disease detection using machine learning.

1.2 Motivation:

I choose this project as I always had an Internet Towards the Agriculture. When I Heard the News Of Plant Disease in our own Land And My Grandfather Who is Struggling , when I see this option I suddenly chose this and by this project's enhancement I clearly depict a view over million's of village farmers problem.

Automating the process of detecting and identifying through visual inspection (cognitive) is the motivation behind this work. This is made possible with the availability of images of the plant or parts of plants, since most diseases are reflected on the leaves.

1.3 Objective:

The Main Objective of this Project is to create a detection system that detects the disease in the plants through scanning the leaves and reducing the crop loses. The other objective is to provide farmers a user friendly tool to detect disease and reducing the efforts required for manual inspection.

Identify diseases

Detect the presence of infections in plants Detect early

Identify diseases as early as possible to prevent the spread of disease

Provide treatments

Provide recommendations for treating the disease

Improve crop yields

Help farmers improve crop yields by providing tools to identify diseases

Reduce labor

Reduce the amount of labor required to manually monitor plant diseases

1.4 Scope of the Project:

As Agriculture is the Backbone of India and the disease in the plants are gone to be playing the vital role in destroying plants. The Disease Detection System in Plants is Never Going To Be Destroyed.

CHAPTER 2

Literature Survey

2.1 Review

The occurrence of plant diseases has a negative impact on agricultural production. If plant diseases are not discovered in time, food insecurity will increase [1]. Disease-infected plants usually show obvious marks or lesions on leaves, stems, flowers, or fruits.

Usually, the leaves of plants are the primary source for identifying plant diseases, and most of the symptoms of diseases may begin to appear on the leaves.

Main reasons of crop diseases are the infections such as insectpests, bacteria, fungi and viruses. These diseases are found and can spread in all parts of the plants like in stem, vegetables, fruits and others can be detected by one of the listed below:

- Discerning the affected area
- Retrieving the features set of the affected area
- Identifying and categorizing the diseases [2]

Grid search and random search are the most popular hyperparameter tuning approaches in deep learning. High-performance computing power is needed to train the deep learning algorithm with better efficiency and less training time [3].

The data augmentation technique improves the diversity of training data without collecting new data. The augmentation techniques are GAN, flipping, cropping, shifting, principal component analysis (PCA), color, noise and rotation. The result shows that the training performance of the cropping, flipping, GAN and rotation are higher than the other augmentation techniques [3].

2.2 Existing models, techniques, or methodologies related to the problem.

For identification of different plant disease, sensors for imaging system are deployed to accumulate the data for study of leaves from different aspect. Various useful imaging techniques include thermal imaging, multispectral imaging, fluorescence imaging, hyper spectral imaging, visible imaging, MRT. Also 3D imaging methods are also tested along various other methods. In next sections we present a state of survey on these techniques along with their applications in different ways [3].

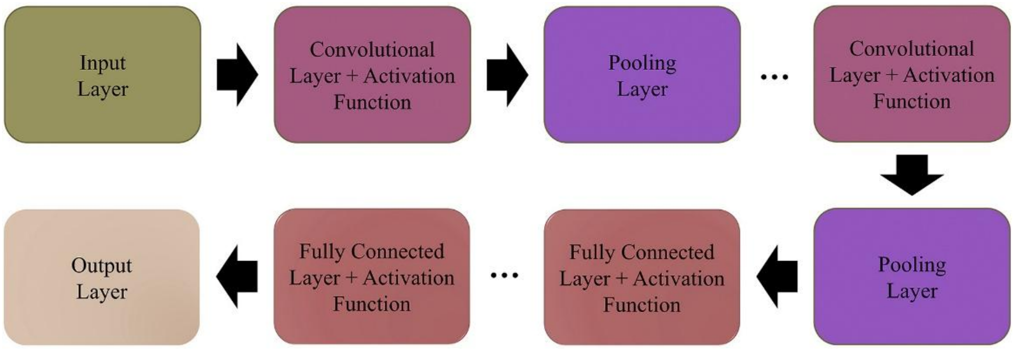

Fig 1 - Depicts the architecture of a typical CNN that contains one Input layer, one Output layer, a set of Convolutionallayers (each with an activation function), Pooling layers, and Fully Connected layers (each with an activation function).

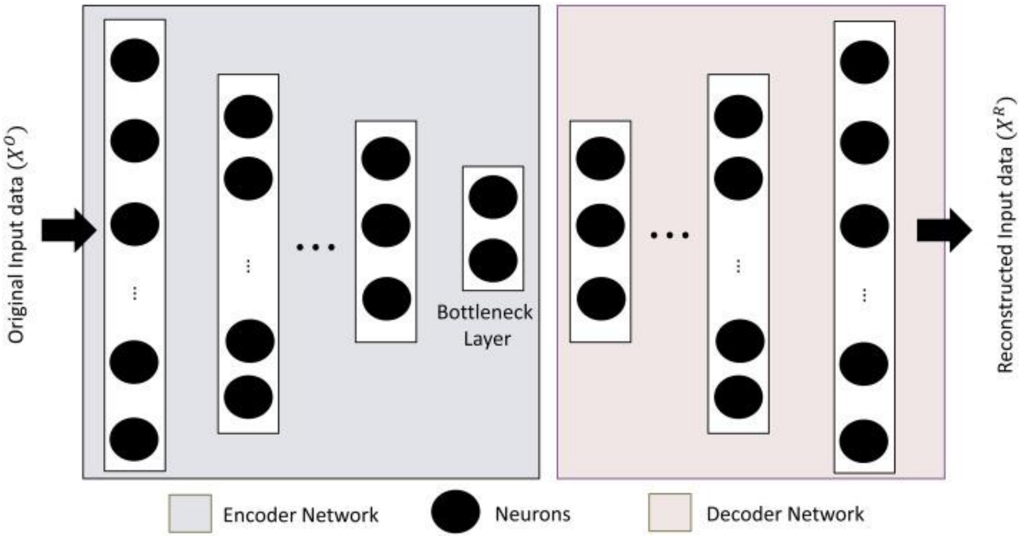

Fig 2 - Architecture of a *N* layer Autoencoder.

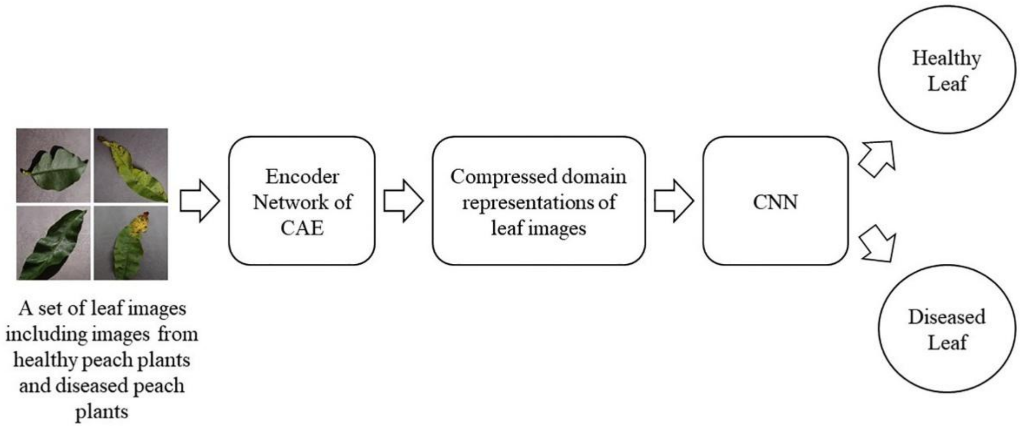

Fig 3 - The block diagram of the proposed hybrid model

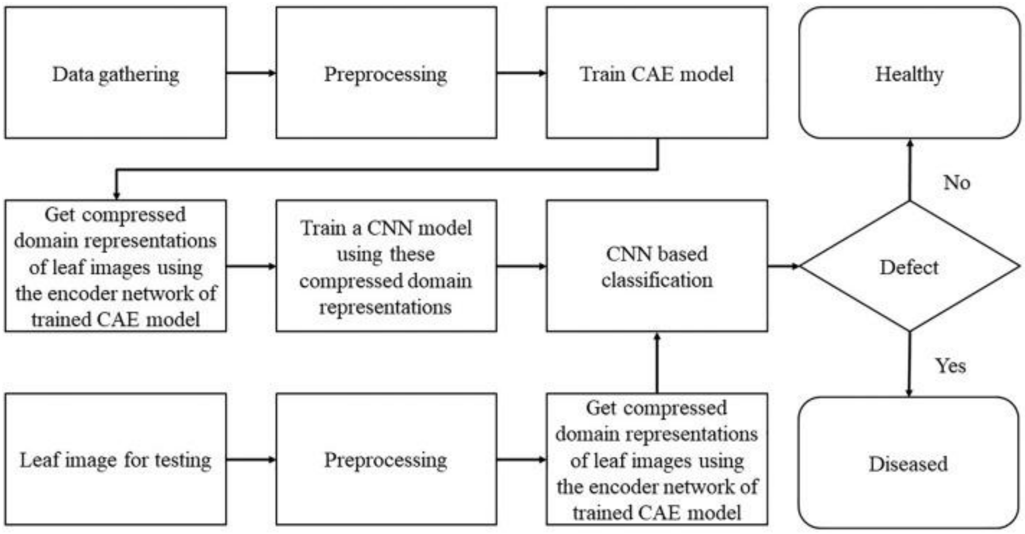

Fig 4 - The flow-diagram to demonstrate the proposed methodology

2.2.1. Magnetic resonance imaging

Also termed as NMR, meaning nuclear magnetic resonance scanner, it is mostly known as magnetic resonance imaging device, is usually identified for its powerful magnets. These magnets are good as they efficiently polarize and further excites the focused proton singly included in water molecules present in the tissue, helping in a detectable signal spatially encoded giving various images of the body. Radio Frequency (RF) pulses are emitted by MRI machines that bind only to oxygen. This system works by initially generating the pulse and transferring it to the examined area of the body. Later they are made to spin in a different orientation by absorption of the send energy. This is called the resonance process involved in the MRI[2].

2.2.2. Photo acoustic imaging

Photo acoustic imaging is a technique that has been derived using hybrid biomedical imaging that is originated around the photo acoustic effect. It involves the amalgamation of different benefits such as optical absorption contrast along ultrasonic spatial resolution involved in deep imaging of diffusive and other regime. The studies bring the fact that photo acoustic imaging can be used for various purposes such as tumor analyzing, mapping of the level of blood oxygen, imaging of the brain activity, and other disease detection, etc[2].

2.2.3. Tomography

Tomography is one of the techniques that involve imaging of a single plane, or an object giving a tomogram. There are different types of tomography such as linear, poly tomography, zonagraphy, computed type or computed axial and Positron Emission type of tomography[2].

2.3Machine Learning (ML) and Deep Learning (DL) Methods

2.3.1. Convolutional Neural Networks (CNNs)

CNNs have been widely used for image-based plant disease detection.

2.3.2. Transfer Learning

Utilizing pre-trained models, such as VGG16 and ResNet50, for plant disease detection.

2.3.3. Support Vector Machines (SVMs)

SVMs have been used for classification of plant diseases based on features extracted from images.

2.4Data-Driven Approaches

2.4.1. Data Mining

Techniques, such as clustering and decision trees, to discover patterns and relationships in plant disease data.

2.4.2. Predictive Modeling

Techniques, such as regression and time series analysis, to predict plant disease outbreaks.

2.5 Gaps And Limitations

Plant disease identification by visual way is more laborious task and at the same time, less accurate and can be done only in limited areas. Whereas if automatic detection technique is used it will take less efforts, less time and become more accurate.

2.5.1 Gaps in Existing Solutions

Accuracy and Reliability

Existing solutions often rely on manual inspection or basic machine learning algorithms, leading to accuracy issues.

Limited Crop Coverage

Many solutions focus on a specific crop or disease, limiting their applicability.

Dependence on Expertise

Some solutions require extensive agricultural knowledge or technical expertise.

High Cost and Complexity

Advanced solutions can be expensive and complex, making them inaccessible to small-scale farmers.

Limited Real-time Capabilities

Existing solutions often lack real-time monitoring and alert systems.

Our project aims to address these gaps by

1. **Developing Advanced AI Algorithms:**

    Utilizing cutting-edge deep learning techniques to improve accuracy and reliability.

2. **Expanding Crop Coverage:**

    Creating a comprehensive database of various crops and diseases.

3. **User-Friendly Interface:**

    Designing an intuitive interface that doesn't require extensive expertise.

4. **Affordable and Accessible:**

    Developing a cost-effective solution that can be accessed by small-scale farmers.

5. **Real-time Monitoring and Alerts:**

    Implementing a real-time monitoring system that sends alerts to farmers.

By addressing these gaps, our project aims to provide a comprehensive, accurate, and accessible plant disease detection solution for farmers worldwide. Crop resistance mechanisms, breeding disease resistant plant varieties, resistance gene identification, cloning and implementation, and pathogens overcoming plant resistance, climate change and abiotic stress on plant resistance, resistance against multiple pathogens.

Important environmental factors that may affect development of plant diseases and determine whether they become epiphytotic include temperature, relative humidity, soil moisture, soil pH, soil type, and soil fertility.

Key challenges include dealing with image quality and object image variations, achieving real-time analysis, and ensuring high-level accuracy across complex conditions.

CHAPTER 3

Proposed Methodology

3.1 System Design

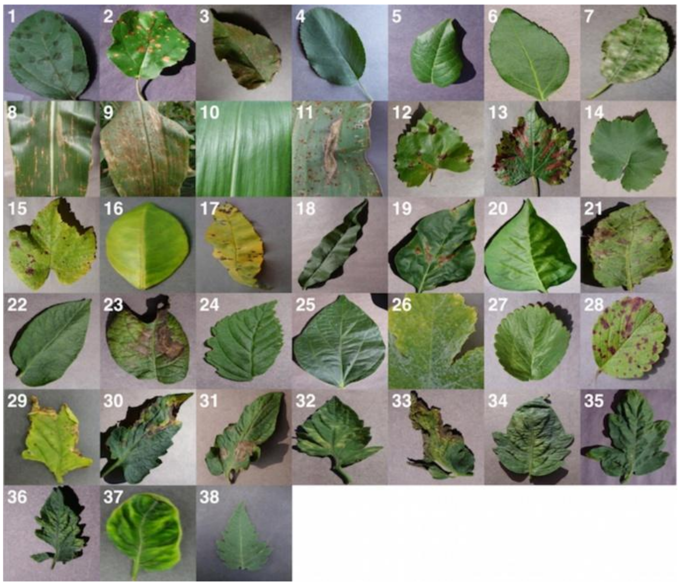

Fig 5 - Example of leaf images from the Plant dataset, representing every crop-disease pair used [5] .

The Above Pictures Clearly Shows Us About The Types of Plant Leaves . Let Us See what it is

(1) Apple Scab, *Venturia inaequalis* **(2)** Apple Black Rot, *Botryosphaeria obtusa* **(3)** Apple Cedar Rust, *Gymnosporangium juniperi-virginianae* **(4)** Apple healthy **(5)** Blueberry healthy **(6)** Cherry healthy **(7)** Cherry Powdery Mildew, *Podoshaera clandestine* **(8)** Corn Gray Leaf Spot, *Cercospora zeae-maydis* **(9)** Corn Common Rust, *Puccinia sorghi* **(10)** Corn healthy **(11)** Corn Northern Leaf Blight, *Exserohilum turcicum* **(12)** Grape Black Rot, *Guignardia bidwellii*, **(13)** Grape Black Measles (Esca), *Phaeomoniella aleophilum*, *Phaeomoniella chlamydospora* **(14)** Grape Healthy **(15)** Grape Leaf Blight, *Pseudocercospora vitis* **(16)** Orange Huanglongbing (Citrus Greening), *Candidatus Liberibacter spp.* **(17)** Peach Bacterial Spot, *Xanthomonas* *campestris* **(18)** Peach healthy **(19)** Bell Pepper Bacterial Spot, *Xanthomonas campestris* **(20)** Bell Pepper healthy **(21)** Potato Early Blight, *Alternaria solani* **(22)** Potato healthy **(23)** Potato Late Blight, *Phytophthora infestans* **(24)** Raspberry healthy **(25)** Soybean healthy **(26)** Squash Powdery Mildew, *Erysiphe cichoracearum* **(27)** Strawberry Healthy **(28)** Strawberry Leaf Scorch, *Diplocarpon earlianum* **(29)** Tomato Bacterial Spot, *Xanthomonas campestris pv. vesicatoria* **(30)** Tomato Early Blight, *Alternaria solani* **(31)** Tomato Late Blight, *Phytophthora infestans* **(32)** Tomato Leaf Mold, *Passalora fulva* **(33)** Tomato Septoria Leaf Spot, *Septoria lycopersici* **(34)** Tomato Two Spotted Spider Mite, *Tetranychus urticae* **(35)** Tomato Target Spot, *Corynespora cassiicola* **(36)** Tomato Mosaic Virus **(37)** Tomato Yellow Leaf Curl Virus **(38)** Tomato healthy.

One of the steps of that processing also allowed us to easily fix color casts, which happened to be very strong in some of the subsets of the dataset, thus removing another potential bias.

This set of experiments was designed to understand if the neural network actually learns the “notion” of plant diseases, or if it is just learning the inherent biases in the dataset.

Figure 2 shows the different versions of the same leaf for a randomly selected set of leaves.

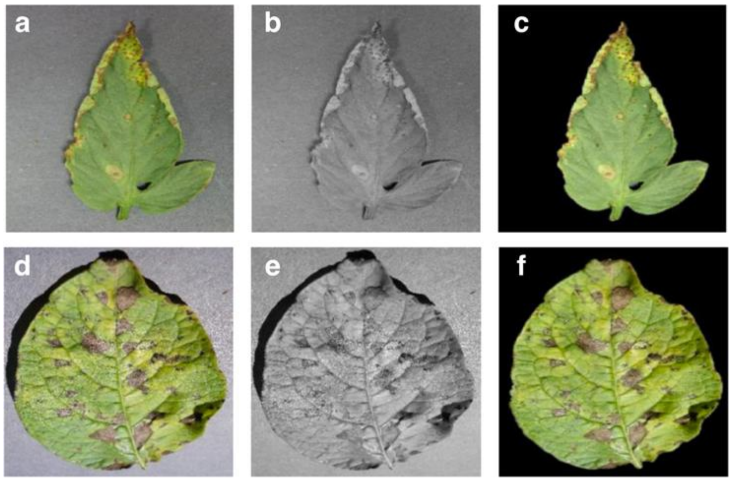

Fig 6 Sample images from the three different versions of the Plant dataset used in various experimental configuration [5]

Every Leaf Represent Each Experimental Configurations like as follows

(A) Leaf 1 color, **(B)** Leaf 1 grayscale, **(C)** Leaf 1 segmented, **(D)** Leaf 2 color, **(E)** Leaf 2 gray-scale, **(F)** Leaf 2 segmented.

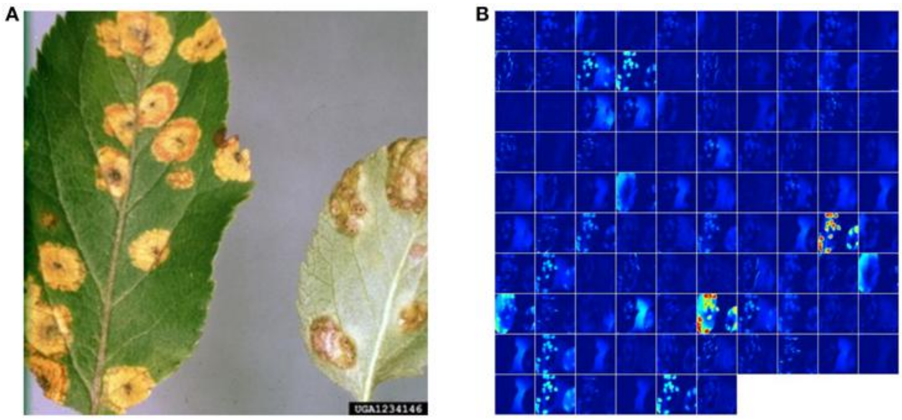

Fig 7 - Visualization of activations in the initial layers of an AlexNet architecture demonstrating that the model has learnt to efficiently activate against the diseased spots on the example leaf.

(A) Example image of a leaf suffering from Apple Cedar Rust, selected from the top-20 images returned by Bing Image search for the keywords "Apple Cedar Rust Leaves" on April 4th, 2016. Image Reference: Clemson University - USDA

Cooperative Extension Slide Series, Bugwood. org. **(B)** Visualization of activations in the first convolution layer(conv1) of an AlexNet architecture trained.

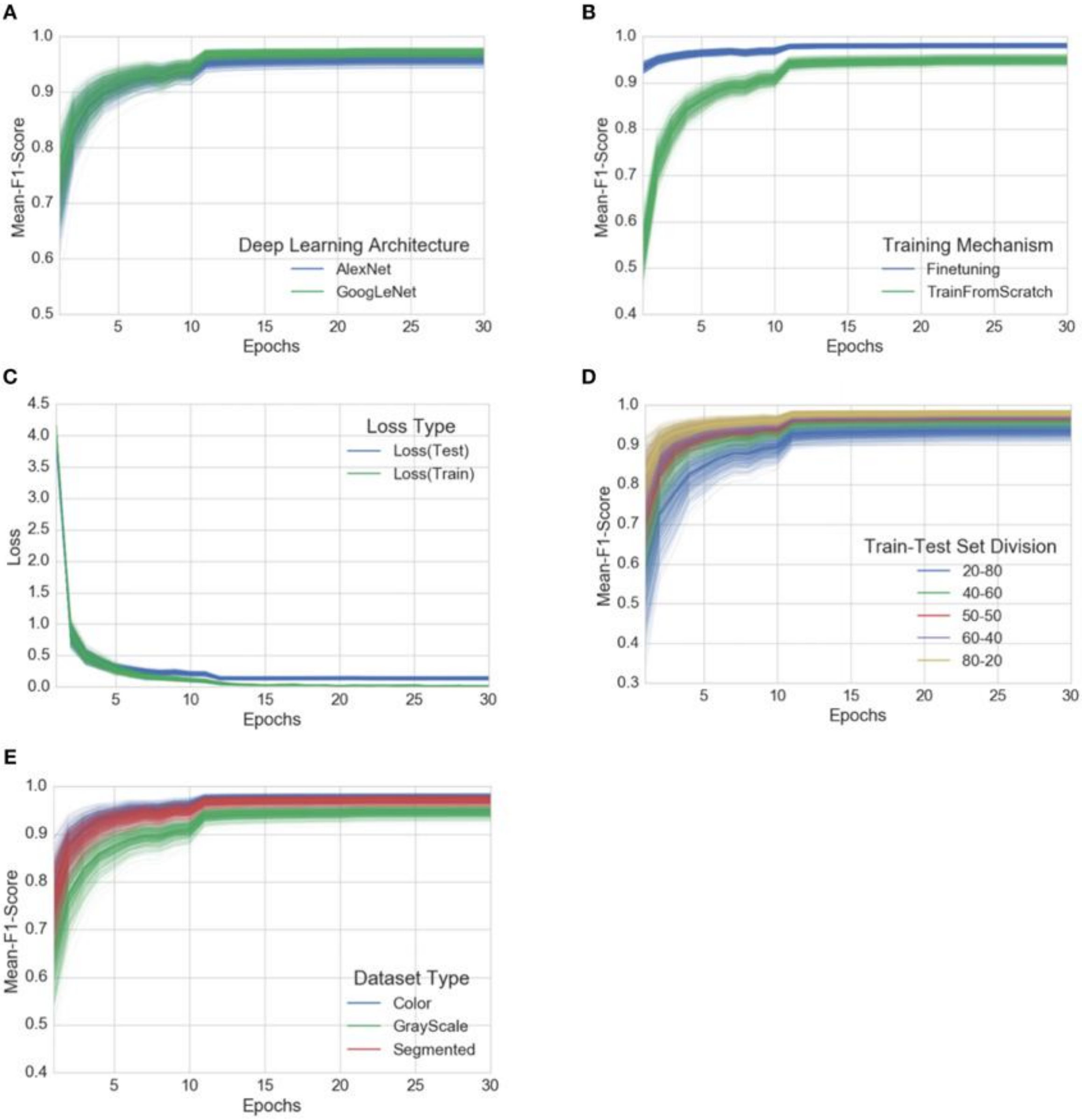

Fig 8 - Progression of mean F_1 score and loss through the training period of 30 epochs across all experiments, grouped by experimental configuration parameters

The intensity of a particular class at any point is proportional to the corresponding uncertainty across all experiments with the particular configurations. **(A)** Comparison of progression of mean F_1 score across all experiments, grouped by deep learning architecture, **(B)** Comparison of progression of mean F_1 score across all experiments, grouped by training mechanism, **(C)** Comparison of progression of train-loss and test-loss across all experiments, **(D)** Comparison of progression of mean F_1 score across all experiments, grouped by train-test set splits, **(E)** Comparison of progression of mean F_1 score across all experiments, grouped by dataset type. A similar plot of all the observations, as it is, across all the experimental configurations can be found in the Supplementary Material.

3.2 Requirement Specification

3.2.1 Hardware Requirements:

Smartphones or Cameras

For farmers to capture images of plants.

IoT Devices

For collecting sensor data (e.g., temperature, humidity, soil moisture).

Drones

For capturing aerial images of crops (optional).

Cloud Servers

For hosting the machine learning model, database, and web application.

GPU Accelerators

For accelerating machine learning model training and deployment.

3.2.2 Software Requirements:

Programming Languages:

- Python for machine learning model development and deployment.
- JavaScript for web application development.
- SQL for database management.

Machine Learning Frameworks:

- TensorFlow or PyTorch for building and training machine learning models.

Deep Learning Libraries:

- Keras for building and training deep learning models.

Computer Vision Libraries:

- OpenCV for image processing and feature extraction.

Web Development Frameworks:

- React or Angular for building the web application.

Database Management Systems:

- MySQL or PostgreSQL for managing the database.

Cloud Platforms:

- Amazon Web Services (AWS) or Google Cloud Platform (GCP) for hosting the cloud servers.

APIs and SDKs:

- APIs for integrating with IoT devices and drones (if used).
- SDKs for integrating with cloud platforms and machine learning frameworks.

Operating Systems:

- Linux or Windows for cloud servers and development environments.

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:

Fig 9 - An example of a healthy and a diseased leaf image of a peach plant

The dataset contains 4457 leaf images of peach plants, which are evenly distributed in two classes: healthy and diseased (Bacterial Spot). The healthy class contains 2160 peach leaf images, and the diseased (Bacterial Spot) class comprise of 2297 leaf images of the peach plant.

Fig 10 - (Top row) original leaf images, (Bottom row) reconstructed leaf images using the convolutional auto encoder network.

Fig 11 – Leaf Of A Apple Plant

Image of a healthy plant leaf with a green color and no visible symptoms of disease.

Detection Result: Healthy

Explanation: This snapshot shows the output of the system when a healthy plant leaf image is uploaded. The system correctly identifies the leaf as healthy and displays the result.

Fig 12 -Leaf of apple plant with yellow color or discoloration

Image of a plant leaf with yellow spots and discoloration.

Detection Result: Infected with Yellow Leaf Spot Disease (confidence level: 85%)

Explanation: This snapshot demonstrates the system's ability to detect plant diseases. In this case, the system correctly identifies the yellow spots and discoloration as symptoms of Yellow Leaf Spot Disease and provides a confidence level of 85%.

Fig 13 - Image of apple plant leaf with powdery mildew symptoms.

Detection Result: Infected with Powdery Mildew Disease (confidence level: 90%)

Recommendation: Apply fungicide treatment and ensure good air circulation around the plant.

Explanation: This snapshot showcases the system's capability to not only detect diseases but also provide recommendations for treatment and management. In this case, the system correctly identifies the powdery mildew symptoms and provides a confidence level of 90%, along with a recommendation for fungicide treatment and improved air circulation.

4.2GitHub Link for Code:

https://github.com/Chandhivisha18/Aicte-2025-Project-Aiml.git

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

Here are some suggestions for improving the model or addressing unresolved issues in future work:

Data-Related Improvements

1. Increase Dataset Size and Diversity: Collect more data from various sources, including different crops, diseases, and environmental conditions.
2. Improve Data Quality: Ensure that the collected data is accurate, complete, and consistent.
3. Use Transfer Learning: Utilize pre-trained models and fine-tune them on the plant disease dataset.

Model-Related Improvements

1. Experiment with Different Architectures: Try various deep learning architectures, such as ResNet, Inception, or DenseNet.
2. Use Ensemble Methods: Combine the predictions of multiple models to improve overall accuracy.
3. Implement Attention Mechanisms: Focus the model's attention on specific parts of the image that are relevant for disease detection.

Real-World Deployment

1. Develop a User-Friendly Interface: Create a simple and intuitive interface for farmers to upload images and receive disease diagnoses.
2. Integrate with IoT Devices: Incorporate data from IoT devices, such as temperature and humidity sensors, to improve disease detection accuracy.
3. Provide Personalized Recommendations: Offer tailored advice to farmers based on the specific disease diagnosis, crop type, and environmental conditions.

Future Research Directions

1. Explore Other Modalities: Investigate the use of other data modalities, such as hyperspectral or multispectral imaging.
2. Develop a Real-Time System: Create a real-time disease detection system that can provide instant feedback to farmers.
3. Investigate Transfer Learning Across Crops: Examine the feasibility of transferring knowledge from one crop to another.

5.2 Conclusion:

Disease detection in plants at the early stages is a hard and challenging task. Many researchers have used different Machine Learning and Deep Learning techniques for automatic plant disease detection. However, most of these techniques either use millions of training parameters or have a low classification accuracy. In this paper, a novel hybrid model was proposed for automatic plant disease detection that was based on two Deep Learning techniques named Convolutional Autoencoder (CAE) network and Convolutional Neural Network (CNN). The proposed hybrid model first obtained compressed domain representations of leaf images using the encoder network of CAE and then used the compressed domain representations for classification using CNN.

Due to dimensionality reduction using CAE, the number of features, and hence the number of training parameters reduced significantly as compared to existing state-of-the-art systems. To test the model, it was applied to detect Bacterial Spot disease in peach plants. The model achieved 99.35% training accuracy and 98.38% testing accuracy by using only 9,914 training parameters. Fewer training parameters used in the proposed hybrid model significantly decreased the time required to train the model for automatic plant disease detection and the time required to identify the disease in plants using the trained model[5].

REFERENCES

[2] Singh, V., Sharma, N., & Singh, S. (2020). A review of imaging techniques for plant disease detection. *Artificial Intelligence in Agriculture*, *4*, 229-242.

[1] Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—a review. *IEEE Access*, *9*, 56683-56698.

[3] Pandian, J. A., Kumar, V. D., Geman, O., Hnatiuc, M., Arif, M., & Kanchanadevi, K. (2022). Plant disease detection using deep convolutional neural network. *Applied Sciences*, *12*(14), 6982.

[4] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, *7*, 1419.

[5] Bedi, P., & Gole, P. (2021). Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network. *Artificial Intelligence in Agriculture*, *5*, 90-101.