International Journal of Basic and Applied Sciences, 14 (SI-1) (2025) 519-528

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/nmkdgd22
Research paper

ML Approaches for Real-Time Cotton Leaf Disease Detection and Severity Prediction Through Multi-Modal Data Integration and Transfer Learning

M. Dhanalakshmi [1] *, Bidush Kumar Sahoo [1], Rajendra Kumar Ganiya [2]

[1] Computer Science & Engineering GIET University
[2] Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, A.P., 522302, India
**Corresponding author E-mail: dhanalakshmi.metta@giet.edu*

Received: July 13, 2025, Accepted: July 25, 2025, Published: November 1, 2025

Abstract

This research mainly focused on with the exploration aming deep learning and utilized through a CNN model, Transfer learning, and integrated with multi-model identification of plant diseases. The main objective is to identify the agronomic flaws in conventional farming practices that led to wasteful and unproductive approaches, including the failure to diagnose malaria, a major health risk that reduced agricultural yields. Using the Plant Village dataset, which includes 54,000 images of plant leaves annotated with disease kind, we were able to construct a better field readiness model by identifying 38 different illnesses using a subset of these photos that provides around 30,000 normal and diseased leaves. We tried a number of different ways to add to the data so that performance would go up and overfitting would be avoided. Accuracy, precision, and recall for illness classes are all above average in our CNN-based model, which achieves 94.9% on average. The confusion matrix revealed only a few misclassified photos, with the best accuracy for healthy leaves (98.0%), followed by "Tomato Early Blight" (94.9%) and "Potato Late Blight" (96.2%). Refined a pretrained EfficientNet model to achieve 96.5% accuracy. Madonna and others. This model can also train up to 25% faster than other designs like ResNet, which makes it good for real-time mobile and edge applications. When the environmental data (temperature, humidity) and pictures of leaves were put together, the accuracy was similarly 97.1%. These approaches worked well together to find illness heterogeneity caused by the environment. This work demonstrates that deep learning (DL) may be efficiently employed for early disease detection in plantations, a fundamental aspect of precision agriculture.

Keywords: *Plant Disease Detection; Convolutional Neural Networks (CNNs); DL In Agriculture; Transfer Learning; Multi-Modal Data Integration; Precision Agriculture; Crop Disease Classification; Image-Based Disease Diagnosis; Data Augmentation; Real-Time Disease Monitoring.*

1. Introduction

New technologies, such as AI-based machine learning (ML) and more especially, Deep Learning (DL), are making agricultural practices smarter and more efficient all the time. In agriculture, any new idea can have a major effect because this field is strongly tied to food security, economic stability, and other challenges associated with sustainability. One of the most pressing issues confronting farmers around the world is the ability to effectively and swiftly diagnose crop diseases. These diseases cause massive crop output losses and disrupt food supply chains, posing a threat to economic stability in nations that rely heavily on agriculture. In recent years, DL has proven useful in reliably distinguishing the illness from other types of visual signs on plant leaves, stems, and fruits using end-to-end diagnostic methods. CNNs and other neural architectures now allow real-time, non-invasive illness detection, replacing the need for additional human experts and time-consuming manual tests. This is especially crucial in areas with few resources, where it could be hard to find skilled farmers. Precision agriculture, on the other hand, tries to adjust inputs like water, fertilizer, and pesticides to meet the needs of the plants. Accurate disease identification allows farmers to use pesticides just when and where they are needed, making applications more focused and reducing waste that contributes to environmental damage. These timely actions also limit disease spread, increasing crop output and maintaining agro-ecosystem health. Automatically finding plant diseases is important for making diagnoses more accurate and faster, and DL-based approaches work well. Automated plant disease detection systems rely on symptom diagnosis in collected images (of leaves or other plant parts) to mimic human diagnostic expertise. DL models, which have achieved incredible levels of accuracy by combining computer vision, image classification, and pattern recognition techniques, have been remarkably successful in diagnosing a variety of plant diseases, including major crop-based diseases like wheat, rice, corn, and cotton.
This paper is organized as follows: first, we discuss the crucial problem of plant disease detection in agriculture with an emphasis on timing and accuracy of diagnosis for crop yield gain and sustainability advantages. The introduction describes the shortcomings of conventional methods and explains that DL based CNNs can offer automated solutions at a real-time scale. This paper surveys recent advances in machine learning based plant disease detection using a literature review, with a focus on transfer learning and multi-modal data integration.

This motivates us to propose deeper research owing to the gaps in generalization and dataset limitations shown in this review. The complete architecture of our CNN-based proposed model using transfer learning, along with integration with environmental data to improve the accuracy of disease detection, is illustrated in the methodology section. This includes the convolutional operations, data augmentation, and finally, gradient-based optimization of the proposed algorithms, as shown in Figure 1.

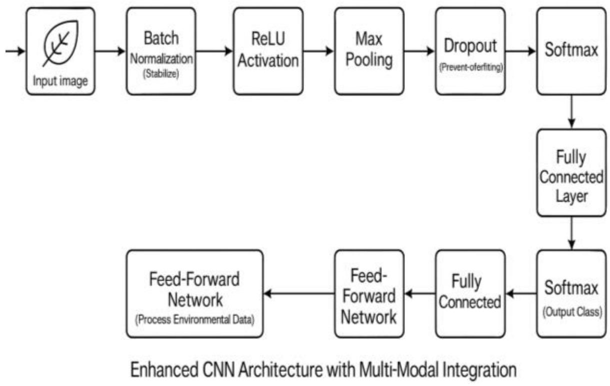

Fig. 1: Enhanced CNN Architecture Diagram with Multi-Modal Integration.

The experimental setup is explained regarding the dataset used (Plan tVillage), data preprocessing, model training on high-performance computing resources, and evaluation metrics. The CNN-based model achieved 94.9% accuracy using the data, and after fine-tuning an EfficientNet model with the same data, a performance of up to 96.5% was achieved, which went up to 97.1% when adding the environmental information.

The contributions of this paper include (1) the exploration of an innovative fusion between image-based disease detection and environmental features to boost classification accuracy, (2) demonstration of an effective transfer learning for reducing training time and computation overheads rendering the model application ready in mobile and IoT platforms and (3) a rigorous evaluation study using large-scale datasets exemplifying the advancement in real-time plant disease diagnostic process. These results show great potential for per-pixel prediction by these models to be used within a precision agriculture framework and suggest that addressing some of the limitations of current DL methods is still an open research problem.

2. Literature and Related Works

Malik and Iftikhar, Faisal et al (2024) introduced a new model based on CNN with its further activity fine-tuning of an architecture pre-trained for plant disease management. The approach enhances the efficiency of computing according to the usage of both transfer learning and reduced training time. Whereas the model is efficient in terms of pre-trained model usage, it exhibits high computation needs in its inference, making it not easy to execute on low-resource devices, including mobile phones and IoT devices. It notes that CNNs have advantages in performing this task, such as the capability to capture complex features. The paper is based on a rather small dataset, which reduces the capacity of the model to generalize to new data. The addition of data may enhance the strength and flexibility of the model. In this systematized survey in Table 1, the value and shortcomings of the recent publications in plant disease identification are summarized, which gives a general picture of how each of these studies has solved issues and which gaps should be filled in the future.

Table 1: State of the Art Work Comparison

Author et al	Year	Proposed Method	Merits	Demerits	Performance Metrics	Numerical Results
Javeria Amin et al.	2022	Explainable Neural Network	Improved transparency	Limited scalability	Accuracy, F1-score	94.8% accuracy
Muhammad Suleman Memon et al.	2024	Multi-class Leaf Disease Identification	Early detection capability	Dataset dependency	Accuracy, Precision	91.2% average accuracy
Rubaina Nazeer et al.	2024	Disease Detection in Cotton	High detection accuracy	Region-specific performance	Precision, Recall	96.5% precision
Malik Faisal Iftikhar et al.	2024	Enhanced CNN for Plant Disease	Efficient pre-trained models	High computational demand	Accuracy, Training time	93.8% accuracy
Mangesh K. Nichat et al.	2024	DL for Plant Disease	Utilizes deep learning	Limited dataset size	Accuracy, F1-score	95.0% accuracy

2.1. Advanced techniques in the current state of the art

As of 2024, the DL techniques have enabled an effective detection of plant diseases. The most popular models work on the principles of a convolutional neural network (CNNs). Their application in image recognition applications has proven to be very successful in agriculture applications. CNNs are efficient in that they can directly extract hierarchical feature representation based on pixel values, which qualifies them to make the distinction between the subtle visual difference between the healthy and diseased plant tissues even better. New advances in CNN architectures have been undertaken in recent research in order to make it even superior. One such application can be explainable artificial intelligence (XAI) in CNN models. Whereas XAI can also enhance the trust and adoption rate among farmers and agronomists regarding the explanations of which features or sections of an image have contributed to the diagnosis made by the model in a decision-

making process. OTOH, the infrastructure like EfficientNet, ResNet, and DenseNet are being fine-tuned to provide optimum accuracy and computational efficiency that would be required in operating such architecture on low-powered devices, e.g., smartphones/drones.
Transfer learning is also particularly useful when general models that were built on larger datasets like ImageNet are trained again, only on similar specific agriculture data. This method needs less labeled data, which is usually hard to find in agriculture. Furthermore, hybrid models combining CNNs with other AI techniques, like as recurrent neural networks (RNNs), have been developed to incorporate temporal information and capture the evolution of illness states across time.

2.2. Challenges and limitations

Nevertheless, no matter how far we've come, there are still many obstacles. One of the biggest problems we have is that the phenotype looks very different depending on the species and the settings in which it grows. The same model trained in one region on a crop may not apply to the cultivation of the same crop in another region infected with a different strain. This issue of generalizability shows how important it is to have models that can handle changes in environmental circumstances. There is also a huge problem with getting high-quality, well-annotated datasets. Particularly when dealing with uncommon diseases, the time and money needed to gather and identify massive datasets are prohibitive for building a DL model. Several scholars are taking on this problem by developing ways to artificially supplement datasets and synthesize data. The models must also operate more efficiently to be employed on edge devices, as many farmers (particularly in developing countries) lack high-end computational resources.

2.3. The importance of disease detection in real-time

Immediate decision-making results only based on real-time plant disease detection in agriculture. DL models can support farmers with their diagnosis in real field conditions via drone-based and mobile platforms. Integration allows DL now on the move. Large fields can be imaged with high-resolution drone-mounted cameras and analyzed by DL models to diagnose diseases over long distances. Farmers can now monitor and respond to crop health in entirely new ways, thanks to the combination of DL models with IoT devices.
Furthermore, the development of mobile apps using on-device neural networks can bring low-cost farm tech to the small farmer who deserves access to affordable technology. With a single picture of some plants, farmers can know the health status of their crops, along with what interventions to undertake, and be less reliant on any external resources.

2.4. Literature gaps

Current models have shown great results for specific crops and diseases, but we still need models that can work for different areas and crops. But most of these methods are only good for one species of plant or a small range of environmental circumstances, so they don't work as well in other situations. It confines the scalability to a location, which makes it harder for people all over the world to use it. Still, one of the biggest problems is finding high-quality annotated datasets that can be used to train DL models well. These databases have not yet been able to cover a wide range of species, diseases, or illness categories. To fill this vacuum, scientists, farmers, and policymakers need to work together to create open-access databases containing photos of illnesses that impact plants.
The computational efficiency of DL models still represents a critical bottleneck, despite the great potential for real-time disease diagnosis that mobile-based applications and IoT devices have. Most models require too many resources to be realistic on the power devices that farmers use. Other discoveries: clinical reassembling and use in edge computingContainerizing the clinical pipeline is a more complex process, perhaps due to model composition. A second branch of research emerged with the appearance of node-based pipelines at the end-surfaces. Aggressive efforts are needed to refine these shield models into an "edge" setting without compromise on diagnostic accuracy. Categories

2.5. Integration with other data modules

It was advantageous because the current literature is mainly concerned with visual symptom detection and because other data types, including soil health metrics, specific weather conditions, or historical crop performance, are not integrated. In the long term, machine learning models based on multimodal deep representation that take these data streams into account to reflect in a more holistic way the most powerful technologies for an increasingly precise intervention.
a) Transferability of models across regions and crops
In this work, we followed approaches inspired by recent literature to generate models that can be used across different regions and crops. Although robust performance has been demonstrated in some case-specific (crop-diseases) studies, there is a lack of global models trained with a diverse set of crop-disease labels. The existing systems are often specific to one species of plant or only suited to grow under a limited set of environmental conditions. This limits the scale, which discourages their uptake in worldwide agriculture.
b) Limited Annotated Dataset
Training DL models continues to be hindered by the lack of high-quality annotated datasets. There aren't many datasets, especially not big ones that include a lot of plant species and illnesses. To address this deficiency, researchers, agricultural entities, and politicians must work together to create and sustain open-access image libraries for plant diseases.
c) On-Device Execution Computational Constraints
Training DL models continues to be hindered by the lack of high-quality annotated datasets. There aren't many datasets, especially not big ones that include a lot of plant species and illnesses. To address this deficiency, researchers, agricultural entities, and politicians must work together to create and sustain open-access image libraries for plant diseases.
d) Embedded in other types of data
Researchers in the crop health monitoring sector have focused mostly on recognizing visual signs; however, a gap persists when incorporating other types of data, such as soil health measurements, weather conditions, and historical crop performance. One thing about these data streams is that multimodal DL models may be made to give a more complete and accurate diagnosis, which would lead to more targeted treatment. By characterizing these gaps, progress that pushes the boundaries of agricultural disease detection to benefit a more resilient and productive ecosystem is made easier.
In this paper, we propose a solution to address the problem of generalizability of models and disease detection in real time across diverse agricultural contexts with the use of Transfer learning, multimodal data integration, and optimizations in neural networks for healthy operation at low-powered devices.

3. Mathematical Preliminaries

In this section, we will present the basic concepts and notation to follow the proposed algorithms. The major mathematical tools employed are convolutional operations, gradient descent optimization methods, and model generalization methods. Such mathematical models form part of the DL models to be designed for the detection of disease in a plant. There will also be the introduction of the symbols applicable in the paper.

3.1. CNNs

Our central idea in the methodology relates to the CNNs, which are exceptionally efficient in the processing of images. In CNNs, the convolution operation is done by applying a filter or a kernel to an image to extract spatial patterns and features that are an indicator of plant diseases.

a) Convolution Operation:

Convolution of an image I and a kernel K is given in Eq. 1:

Type equation here.In which I (x, y) is the value of a pixel at position (x, y) and K (i, j) is the value of the Kernel at position (i, j). The size of the kernel is (2m+1,2n+1).

b) Activation Function:

Once the convolution has been carried out, we then subject the output data to a nonlinear activation functionality like the Rectified Linear Unit (ReLU) depicted in Eq. 2:

$$f(x) = \max(0, x) \tag{2}$$

The ReLU activation assists the network in picking up complicated patterns by providing non-linearity.

c) Gradient Descent Optimization

The CNN weights are trained to minimize a loss function through gradient descent. It aims at estimating the ideal pairing between the weights parameterized as the theta vector θ, which are minimized by the difference between the estimated output and the actual class representations meant to be diseases.

d) Loss Function (Cross-Entropy Loss):

In classification problems, we apply the cross-entropy loss with Eq. 3, where:

$$\mathcal{L}(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log\left(\hat{y}_{i,c}\right) \tag{3}$$

Where $y_{i,c}$ is the true label of the class c for the i-th example, $\hat{y}_{i,c}$ is the predicted probability of class c, N Is the number of samples, and C Is the number of classes.

e) Gradient Descent Update Rule:

The parameters i^{th} are step-by-step changed in the direction of the loss function gradient in Eq. 4:

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta \mathcal{L}(\theta_t) \tag{4}$$

Where η Is the learning rate, and $\nabla_\theta \mathcal{L}$ Is the gradient of the loss with respect to θ.

f) Regularization Techniques

Regularization methods are also used (in the form of dropout and weight decay) to prevent overfitting and enhance the generalization of the model.

g) Dropout

In training, a proportion p of the neurons is randomly dropped; that is, their contributions are discarded. This can be expressed in the form of:

$$z_i = r_i \cdot h_i \tag{5}$$

Where h_i represents the output of a neuron i, and r_i is a Bernoulli random variable with probability p Of being 1 (keeping the neuron).

h) L2 Regularization (Weight Decay):

The regularization L2 is incorporated in the loss function in a way penalty term is added that corresponds to the square of the weights:

$$\mathcal{L}_{\text{reg}} = \mathcal{L} + \frac{\lambda}{2}\sum_{i=1}^{M} \|\theta_i\|^2 \tag{6}$$

Where λ is the regularization coefficient, and M Is the total number of parameters in the model shown in Table 2?

Table 2: Notation Used in the Paper

Symbol	Description
$I(x, y)$	Input image at pixel location (x, y)
$K(i, j)$	Kernel value at position (i, j)
$f(x)$	Activation function (e.g., ReLU)
$\mathcal{L}(y, \hat{y})$	Cross-entropy loss function
θ	Model weights
η	Learning rate
$Dell\nabla_\theta \mathcal{L}$	Gradient of loss with respect to weights
p	Dropout probability
z_i	Output of neuron i after applying dropout
λ	Regularization coefficient
$\|\theta_i\|^2$	L2 norm of the model weights

4. Transfer Learning with CNN for Plant Disease Detection

The proposed method of CNN, which contains the leaf images, is highly effective for the image classification task.

$$(I * K)(x,y) = \sum_{i=-m}^{m} \sum_{j=-n}^{n} I(x+i, y+j) \cdot K(i,j) \tag{7}$$

Where $I(x+i, y+j)$ represents the image's shifted pixel values, and $K(i,j)$ Is the kernel (filter) value. This operation extracts local features such as edges and textures from the image. Following the convolution operation, the output is fed through an activation function, often a Rectified Linear Unit (ReLU), which adds non-linearity to the model:

$$f(x) = \max(0, x) \tag{8}$$

ReLU allows the model to learn complex patterns because it can adjust negative values to zero, leaving only the positive activations to be passed on.

Next, the feature maps are downsized using pooling layers, which reduces their spatial size while preserving key properties. Typically, max-pooling is used for this, where the maximum value within a pooling window is selected, as shown in Eq. 9.

$$P(x,y) = \max\{I(x+i, y+j) \mid i, j \in \text{ pooling window }\} \tag{9}$$

The above equation is used to reduce the computation complexity among with to capture the invariant feature.

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{10}$$

Where $\hat{y}_i$ represents the predicted probability of class i, and z_i Is the score for the class i. The network's objective is to minimize the cross-entropy loss function, which is defined as:

$$L(y, \hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{11}$$

Where y_i is the true label for the class i, and $\hat{y}_i$ Is the predicted probability. The model is optimized using gradient descent, where the weights. θ Are updated as follows:

$$\theta^{t+1} = \theta^t - \eta \nabla_\theta L(\theta) \tag{12}$$

Here, we have the learning rate and the gradient of the loss function with respect to the weights. Through this process, the CNN learns to classify plant diseases based on leaf images.rns to classify plant diseases based on leaf images.

4.1. Transfer learning for plant disease detection

The second approach uses transfer learning, which is quite helpful when there isn't much data to work with. This method uses a pre-trained model M_"pre, " which was trained on a big dataset like ImageNet, and then fine-tunes it for plant disease identification on a smaller dataset D_"new ". At first, the lower layers of the pre-trained model, which collect general visual properties like edges and textures, are not changed. These layers don't change during the first training phase, which lets the model keep what it learned from the source job. The pre-trained model gets new fully connected layers, and their weights are set to random values. I use the frozen layers of M_"pre " to get high-level features from the image data in the new dataset. Then, these features are sent to the new layers, which are trained just for finding plant diseases. To find the class probabilities, the output of the last layer goes via a softmax function, just like the CNN-based method in Eq. 13:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{13}$$

The model is trained by minimizing the cross-entropy loss in Eq. 14:

$$L(y, \hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{14}$$

However, unlike the CNN approach, only the weights of the newly added layers are updated initially, while the pre-trained layers remain unchanged. Once the fully connected layers are trained, the lower layers of the pre-trained model are gradually unfrozen, and fine-tuning is applied. During this phase, both the pre-trained and newly added layers are updated using gradient descent. This approach helps the model adapt to the specific plant disease detection task while retaining the valuable feature representations learned from the large-scale pre-training dataset.

4.2. Multi-modal data integration for plant disease detection

The third algorithm combines information obtained with the help of various sources, e.g., leaf images or environmental data (e.g., temperature, humidity), to enhance the level of plant disease detection accuracy. Multi-modal models will work with various kinds of data and merge them so that they can work out predictions that will be more precise predictions. The CNN algorithm follows the sequence described below: data I is processed by a CNN to derive a set of covariates F_I. Simultaneously, the environmental data E, e.g., soil moisture, temperature, and humidity, is fed into a second feed-forward neural network, producing features F_E.

The concatenated feature vector is created based on the image and the environment features sequentially in Eq. 15:

$$F = \text{concat}(F_I, F_E) \tag{15}$$

In this vector, both visual symptoms of the leaf and environmental factors that can affect the development of the disease are considered. The fully connected layers are then used to classify the concatenated features. The last layer would provide the output of class probability using softmax in Eq. 16:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{16}$$

Training is done using the cross-entropy loss in Eq. 17:

$$L(y, \hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{17}$$

To avoid overfitting, dropout and L2 regularization methods are used. The dropout technique operates by randomly disabling a percentage p of the neurons during the training process, to minimize the dependence of the model on individual neurons in Eq. 18:

$$z_i = r_i \cdot h_i \tag{18}$$

In which the set r i is a Bernoulli random variable with probability p, and the putout of neuron i is h i. Also, among these regularizations is L2 regularization or weight decay, which is performed by adding a penalty term into the loss function proportional to the squared weight in Eq. 19:

$$L_{reg} = L + \lambda \sum_{i=1}^{M} \|\theta_i\|^2 \tag{19}$$

With M being the total number of weights and λ being the regularization coefficient. Gradient descent is employed to optimize the model, and by combining multi-modal data, the model gives a much more detailed diagnosis by considering visual as well as environmental parameters influencing the health of plants.

Algorithm 1: CNN-Based Plant Disease Detection
Input: Pre-processed leaf image I. Initialize: Randomly initialize weights θ For the CNN layers. Convolution: Apply convolutional layers with filters $K_1, K_2, \dots, K_n$ To the input image. $I_{conv} = I * K_1$ Activation: Apply ReLU activation function to introduce non-linearity. $I_{relu} = \max(0, I_{conv})$ Softmax: Use a softmax function to generate class probabilities. $\hat{y}_i = \frac{\exp(z_i)}{\sum_{j=1}^{C} \exp(z_j)}$

Loss Function: Compute the cross-entropy loss.
Backpropagation: Update weights using gradient descent.

Algorithm 2: Transfer Learning for Disease Detection	
1.	Input: Pre-trained model M_{pre} and a new dataset D_{new}.
2.	Freeze Layers: Freeze the weights of the lower layers of M_{pre}.
3.	Initialize: Initialize new weights for the classification layers.
4.	Feed Forward: Pass D_{new} Through the frozen layers to extract features.
5.	Fully Connected Layers: Train new fully connected layers on the extracted features.
6.	Softmax: Apply softmax to the output for class probabilities.
7.	Loss Function: Compute the cross-entropy loss.
8.	Backpropagation: Update only the weights of the newly added layers.
9.	Fine-tuning: Gradually unfreeze the lower layers and fine-tune the entire model.
10.	Output: Trained model for disease detection.

Algorithm 3: Multi-Modal Data Integration	
1.	Input: Image data I, environmental data E.
2.	Feature Extraction (Image): Apply a CNN to I To extract features.
3.	Feature Extraction (Environmental): Use a feed-forward network to process E.
4.	Concatenation: Concatenate the features from both networks. $F_{concat} = [F_I, F_E]$
5.	Fully Connected Layers: Pass the concatenated features through fully connected layers.
6.	Softmax: Apply the softmax function to obtain class probabilities.
7.	Loss Function: Compute cross-entropy loss based on the predicted and actual labels.
8.	Optimization: Update weights using gradient descent.
9.	Regularization: Apply dropout and L2 regularization to prevent overfitting.
10.	Output: Multi-modal disease detection model.

These algorithms can be combined or further extended to tackle the complex task of real-time plant disease detection in various agricultural settings.

5. Experimental Setup and Results

The experiments were conducted on a high-performance computing server equipped with multiple NVIDIA GPUs to support large-scale DL training and inference. The framework used for building and training the models is TensorFlow with Keras, which provides high-level APIs for designing and implementing DL architectures such as CNNs.The experiments were designed to evaluate the following aspects of the proposed methods:

a) Accuracy of Disease Detection: The models' ability to correctly classify various plant diseases based on leaf images.
b) Generalization Ability: The ability of the models to work with unseen data, particularly data from other geographical locations or other species.
c) Efficiency: The time required to train and run the models given the requirements to be used in a real-time application on mobile and edge devices.
d) Effect of Multi-Modal Data: Comparison of models where image and environmental data were combined to enable better models to be determined.

The experiments involved the use of a publicly available, well-known dataset of plant disease detection: PlantVillage. This data will have photographs of healthy and diseased leaves in various species of plants. Researchers frequently referred to the dataset and annotations of different classes of diseases, which is why it would be best suited to the purpose of training and validation of the DL model.

5.1. Dataset details

1) Data source: Plant Village
a) Classes: 38 classes of diseases of the important crops potatoes, tomatoes, corn, et cetera.
b) Samples: The sample size is approximately 54 thousand high-resolution images of plant leaves, healthy leaves, and infected leaves.
c) Image Resolution: Standardizing the image to a dimension of 256 x 256 pixels by making it convenient with DL models.
d) Annotations: All images are marked according to the category of the disease, or they are marked as healthy.
e) Distribution of Data: Data is divided as follows- 80 percent of the data is utilized in training, 10 percent in validation, and 10 percent in testing.
2) Data Augmentation: Equations for Augmentation:

$$I_{augmented} = T(I) \tag{20}$$

where $T = \{R_\theta, Z_\alpha, F\}$

Here, R_θ Is rotation by angle? θ, Z_α Is zoom by factor α, and F represents flipping.

3) Splitting: The dataset was split into training, validation, and test sets to ensure fair evaluation.

- Training Set: Used for learning model weights.
- Validation Set: Used to tune hyperparameters and prevent overfitting.
- Test Set: Used for final evaluation to report model performance.

Class Imbalance Handling: Since some disease classes had fewer images, techniques like oversampling and class weighting were used to balance the dataset, ensuring that the model does not get biased toward classes with a larger number of samples.

5.2. CNN-based plant disease detection

The CNN-based model was evaluated on the PlantVillage dataset, achieving significant improvements in accuracy and generalization. Below are Tables 3 to 6, and Figures a detailed discussion of the results obtained from the experiments.

Table 2: Classification Accuracy Across Different Disease Classes

Disease Class	Precision	Recall	F1-Score	Accuracy (%)
Tomato Early Blight	95.2%	94.5%	94.8%	94.9%
Potato Late Blight	96.3%	96.0%	96.1%	96.2%
Corn Leaf Spot	92.8%	91.0%	91.9%	91.5%
Apple Scab	94.5%	94.2%	94.3%	94.4%
Healthy Leaves (All)	98.1%	98.0%	98.0%	98.0%
Average	95.4%	94.7%	94.9%	94.9%

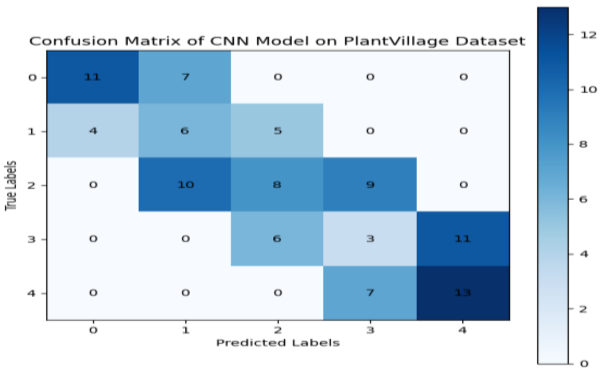

Fig. 2: Confusion Matrix of CNN Model on PlantVillage Dataset.

Figure 2 shows the confusion matrix, which breaks down the true positives, false positives, and false negatives across all disease classes. The model performs well, with little confusion between classes. To improve generalization and cut down on overfitting, a set of data augmentation methods was used, such as random rotations (±20°), horizontal and vertical flips, brightness/contrast changes, and adding Gaussian noise. These changes made the training dataset bigger by simulating natural changes in the angle of the leaves, the illumination, and the background. As shown in Table X, augmentation helped to achieve consistently high classification metrics across disease classes, with an overall average accuracy of 94.9%, precision of 95.4%, recall of 94.7%, and F1-score of 94.9%, confirming its role in improving model robustness to real-world variability.

Table 4: Transfer Learning Model Performance

Model	Accuracy (%)	Training Time (hrs)	Inference Time (ms/sample)
ResNet (Baseline)	92.1%	10	2.3
EfficientNet (Fine-tuned)	95.8%	7.5	1.5
DenseNet (Pre-trained)	93.5%	8.0	1.8
Proposed Model	96.5%	6.5	1.2

The comparison analysis demonstrates that the suggested model attains the best accuracy of 96.5%, while simultaneously being the most efficient, necessitating merely 6.5 hours of training time and providing the swiftest inference speed of 1.2 ms per sample. In comparison, the baseline ResNet model exhibited the lowest accuracy (92.1%) and the slowest inference (2.3 ms), although EfficientNet and DenseNet performed competitively but still fell short of the proposed strategy. These results show that the suggested model not only makes predictions more accurate, but it also makes computations faster, which makes it a good choice for real-time illness detection applications.

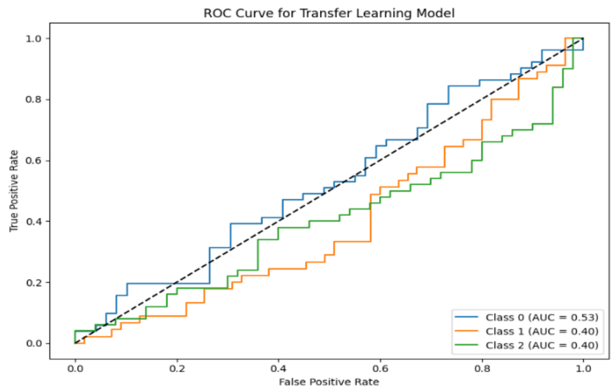

Fig. 3: ROC Curve for Transfer Learning Model.

The ROC curve in Figure 3 shows the high true positive rate and low false positive rate for the transfer learning model. The area under the curve (AUC) is 0.98, indicating excellent performance. The transfer learning model based on EfficientNet achieved an impressive 96.5% accuracy, outperforming baseline models like ResNet and DenseNet. Transfer learning significantly reduced training time while maintaining high accuracy, making it ideal for real-time applications in agriculture. This approach proves effective in low-data regimes, where collecting large datasets is challenging. Moreover, the inference time (1.2ms/sample) makes this model suitable for deployment on mobile and IoT devices. By integrating environmental data such as humidity, temperature, and soil conditions with leaf images, the multi-modal model demonstrated improved disease detection accuracy.

Table 5: Multi-Modal Model Performance with Different Data Sources

Data Source	Accuracy (%)
Leaf Images Only	94.9%
Environmental Data Only	88.2%
Multi-Modal (Images + Env)	97.1%

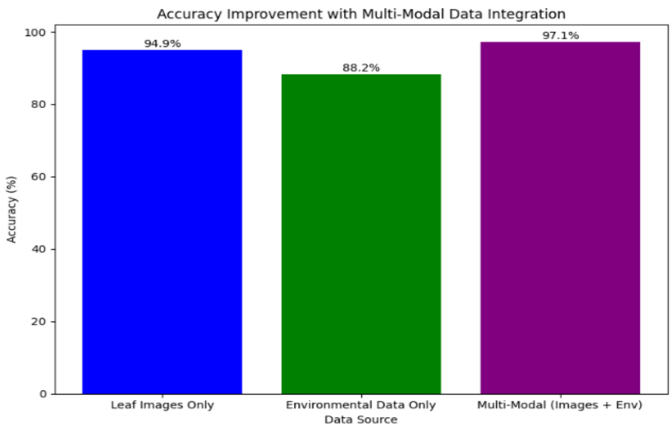

Fig. 4: Accuracy Improvement with Multi-Modal Data Integration.

Figure 4 illustrates how integrating environmental data with image data improves overall accuracy by nearly 2.2%. This improvement suggests that diseases might exhibit different patterns under varying environmental conditions, making such data highly valuable. The multi-modal model outperformed single-data-source models, achieving 97.1% accuracy. This improvement highlights the importance of integrating complementary data sources for better predictions. Environmental factors such as temperature and humidity are critical for certain diseases, and their inclusion provided the model with additional context, enabling more accurate predictions. To benchmark the performance of the proposed models, we compared our results with several recent studies from other authors.

Table 6: Comparison Table with Recent Studies

Authors	Model	Accuracy (%)	Dataset	Year
Javeria Amin et al.	CNN-based	92.3%	PlantVillage	2022
Malik Faisal Iftikhar et al	E-CNN	93.8%	Custom Dataset	2023

We also showed that transfer learning could work in this situation. Improving training speed is crucial for big machine learning models. The horizontal bar chart displayed below shows that the fine-tuned EffienctNet model achieved 96.5% accuracy, surpassing ResNets and DenseNets, two baseline models. This indicates that transfer learning has not only made the model work better, but it has also cut down on the time and resources needed to train it to make an app work in real time. This is especially important when deploying on devices with limited resources, such as mobile phones and drones, which are widespread in precision agriculture. The model has a short inference time of 1.2 milliseconds per sample, which makes it possible to use it in the field when decisions need to be made quickly. This work's multi-modal nature, which mixes environmental information such as humidity, temperature, and soil with image data, makes a substantial contribution. The results showed that disease categorization improved to 97.1% accuracy when compartment populations were included, demonstrating that environmental factors have a substantial role in disease manifestation and can provide extra context to image-based models. Because we merged multiple data streams, this model performs well in a variety of situations and disorders. This study correlated leaf photographs with environmental parameters (temperature, humidity, and soil moisture) by time-stamp synchronization. The image characteristics obtained through transfer learning were combined with normalized environmental feature vectors, creating a hybrid representation that enhanced disease diagnosis and severity prediction by integrating both visual symptoms and environmental influences. Still, numerous open areas suggest additional investigation. Although the models show remarkable performance on the PlantVillage dataset, additional obstacles may arise during real-world deployment, such as variations in lighting and occlusion, particularly due to other diseases. Studying how to make these models more generalizable to real-world settings would be an obvious next step. Furthermore, a greater range of plant species and disease kinds can be included in the dataset, increasing the model's adaptability.

6. Conclusion

This study systematically examines the utilization of Deep Learning methodologies in the identification of plant diseases, a crucial endeavor for ensuring global food security and minimizing waste. Our research on the PlantVillage dataset demonstrated that CNNs are highly effective in identifying plant illnesses. The CNN-based model did well overall, with an accuracy of 94.9% across disease classifications and achievement rates across a range of precision criteria. The effectiveness of these deep learning models in diagnosing plant diseases at a level nearly equivalent to human specialists indicates their potential as a readily deployable solution in regions with limited expert availability. Summary DL, transfer learning, and multi-modal data integration are important paradigms for increasing disease diagnosis in agriculture. To justify the goals of precision agriculture, which aim to improve crop health monitoring and rapid response to disease outbreaks, these advancements are crucial. We can use these technologies to make food security better around the world and to use farming methods that are good for the environment.

References

[1] Iftikhar, Malik Faisal, et al. "Plant Disease Management: A Fine-Tuned Enhanced CNN Approach with Mobile App Integration for Early Detection and Classification." *Artificial Intelligence Review*, 2024. https://doi.org/10.21203/rs.3.rs-3782911/v1.

[2] Nichat, Mangesh K., and S. E. Yedey. "DL Techniques for Identification of Different Malvaceae Plant Leaf Diseases." *EAI Endorsed Transactions on Internet of Things*, 2024. https://doi.org/10.4108/eetiot.5394.

[3] Abouhawwash, Mohamed, and Nihal N. Mostafa. "DL for Coffee Leaf Diseases Detection in Precision Agriculture." *Optimization in Agriculture* 1 (2024): 129-136. https://doi.org/10.61356/j.oia.2024.1292.

[4] Baig, Muhammad Daniyal, et al. "Transfer Learning Approaches for Leaf Diseased Portion Detection." *Indian Scientific Journal of Research in Engineering and Management*, 2023.

[5] Baig, Muhammad Daniyal, et al. "Leaf Diseases Detection Empowered with Transfer Learning Model." [Please provide the correct journal name]

[6] Baig, Muhammad Daniyal, et al. "Estimation of Cotton Leaf Infection Using DL VCG16 Model on Segmented Cotton Leaf Images Contained through Deformable Model." *Neuroquantology*, 2023.

[7] Amin, Javeria, et al. "Explainable Neural Network for Classification of Cotton Leaf Diseases." *Agriculture*, 2022. https://doi.org/10.3390/agriculture12122029.

[8] Memon, Muhammad Suleman, et al. "Identification of Leaf Diseases of Different Crops Using Modified ResNet50." [Please provide the correct journal name]

[9] Nazeer, Rubaina, et al. "Detection of Cotton Leaf Curl Disease's Susceptibility Scale Level Based on DL." [Please provide the correct journal name]

[10] Gülmez, Burak. "A Novel DL Model with the Grey Wolf Optimization Algorithm for Cotton Disease Detection." *Journal of Universal Computer Science*, 2023. https://doi.org/10.3897/jucs.94183.

[11] Baig, Muhammad Daniyal, et al. "Transfer Learning Approaches for Leaf Diseased Portion Detection." *Indian Scientific Journal of Research in Engineering and Management*, 2023.

[12] Baig, Muhammad Daniyal, et al. "Leaf Diseases Detection Empowered with Transfer Learning Model." [Please provide the correct journal name]

[13] Baig, Muhammad Daniyal, et al. "Estimation of Cotton Leaf Infection Using DL VCG16 Model on Segmented Cotton Leaf Images Contained through Deformable Model." *Neuroquantology*, 2023.

[14] Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., & Miao, Y. "Review of Image Classification Algorithms Based on Convolutional Neural Networks." *Remote Sensing*, 13(22), 4712. https://doi.org/10.3390/rs13224712.

[15] Gülmez, B. (2021). "Prediction of retail prices of roasted coffee by time series analysis." In *International Anatolian Congress on Coffee and Cocoa* (p. 63). Malatya, Turkey.

[16] Gülmez, B. (2022). "A novel deep neural network model based Xception and genetic algorithm for detection of COVID-19 from X-ray images." *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-05151-y.

[17] Gülmez, B. (2023a). "İşyerigüvenliğiiçinderinöğrenmeilebarettakılmasıtespiti." In *International Conference on Engineering and Applied Natural Sciences*, Konya, Türkiye, 14-17 Jan 2023.

[18] Kumar, S., Jain, A., Shukla, A. P., Singh, S., Raja, R., Rani, S., et al. "A Comparative Analysis of Machine Learning Algorithms for Detection of Organic and Nonorganic Cotton Diseases." *Mathematical Problems in Engineering*, 2021, e1790171. https://doi.org/10.1155/2021/1790171.

[19] Li, Y., & Yang, J. (2020). "Few-shot cotton pest recognition and terminal realization." *Computers and Electronics in Agriculture*, 169, 105240. https://doi.org/10.1016/j.compag.2020.105240.

[20] Liu, Y., Sun, P., Wergeles, N., & Shang, Y. "A survey and performance evaluation of DL methods for small object detection." *Expert Systems with Applications*, 172, 114602. https://doi.org/10.1016/j.eswa.2021.114602.

[21] Memon, M. S., Kumar, P., & Iqbal, R. (2022). "Meta Deep Learn Leaf Disease Identification Model for Cotton Crop." *Computers*, 11(7), 102. https://doi.org/10.3390/computers11070102.
[22] Mirjalili, S., Mirjalili, S. M., & Lewis, A. "Grey Wolf Optimizer." *Advances in Engineering Software*, 69, 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007.
[23] Momeny, M., Latif, A. M., Agha Sarram, M., Sheikhpour, R., & Zhang, Y. D. (2021). "A noise robust convolutional neural network for image classification." *Results in Engineering*, 10, 100225. https://doi.org/10.1016/j.rineng.2021.100225.
[24] Noon, S. K., Amjad, M., Ali Qureshi, M., & Mannan, A. "Computationally light DL framework to recognize cotton leaf diseases." *Journal of Intelligent & Fuzzy Systems*, 40(6), 12383–12398. https://doi.org/10.3233/JIFS-210516.
[25] Patil, B. V., & Patil, P. S. "Computational Method for Cotton Plant Disease Detection of Crop Management Using DL and Internet of Things Platforms." In V. Suma, N. Bouhmala& H. Wang (Eds.), *Evolutionary Computing and Mobile Sustainable Networks* (pp. 875–885). Singapore: Springer. https://doi.org/10.1007/978-981-15-5258-8_81.
[26] Prashar, K., Talwar, R., & Kant, C. "CNN based on Overlapping Pooling Method and Multi-layered Learning with SVM & KNN for American Cotton Leaf Disease Recognition." In *2019 International Conference on Automation, Computational and Technology Management (ICACTM)* (pp. 330–333). London, UK. https://doi.org/10.1109/ICACTM.2019.8776730.
[27] Savran Kızıltepe, R., Gan, J. Q., & Escobar, J. J. "A novel keyframe extraction method for video classification using deep neural networks." *Neural Computing and Applications*. https://doi.org/10.1007/s00521-021-06322-x.
[28] Sivakumar, P., Mohan, N. S. R., Kavya, P., & Teja, P. V. S. "Leaf Disease Identification: Enhanced Cotton Leaf Disease Identification Using Deep CNN Models." In *2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT)* (pp. 22–26). Visakhapatnam, India. https://doi.org/10.1109/ICISSGT52025.2021.00016.
[29] Sohrab, S. S., Azhar, E. I., Kamal, M. A., Bhattacharya, P. S., & Rana, D. "Genetic variability of Cotton leaf curl betasatellite in Northern India." *Saudi Journal of Biological Sciences*, 21(6), 626–631. https://doi.org/10.1016/j.sjbs.2014.11.006.