International Journal of Applied Mathematics

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

A DEEP LEARNING-BASED FRAMEWORK FOR DISEASE DETECTION AND YIELD ENHANCEMENT IN FIVE MAJOR CROPS

[1]Rikendra, [1]Monika Sharma and [2]Sansar Singh Chauhan

[1] -Department of Computer Science, Amity University Noida, UP, India

[2]-Department of Computer Science, GL Bajaj Institute of Technology and Management Greater Noida, UP, India

Abstract

Precision agriculture is using more and more powerful computer methods to improve crop monitoring and produce management. This study introduces a new deep learning-based framework that combines disease detection in several crops with yield prediction. Its goal is to help with real-time agronomic decision-making. Using the publicly available "Top Agriculture Crop Disease India" dataset, a Hybrid CNN + LSTM + Attention model was made to classify diseases in five main crops: maize, potatoes, rice, sugarcane, and wheat. The suggested Hybrid model had an overall average classification accuracy of 86.9%, which was a statistically significant improvement above baseline CNN designs like ResNet50, EfficientNetB0, and VGG16. We used a CNN + XGBoost pipeline to estimate yield. It used deep features, illness severity scores, and environmental data that CNN extracted. This method gave an average validation accuracy of 93.84% over all five crops. The examination of feature importance indicated that deep features and illness severity are the main factors that affect yield results. The integrated system can make decisions in real time, with a processing time of less than 100 ms per sample. It is set up through a user-friendly online interface. Farmers may use the system to get useful information about how to manage diseases and get the most out of their crops. The results show that the framework could improve precision agriculture methods and make it easier to use with edge computing and multimodal data sources in the future.

Keywords: *Precision Agriculture, Crop Disease Detection, Yield Prediction, Hybrid CNN + LSTM, CNN + XGBoost, Deep Learning, Decision Support System; Real-Time Inference, Multi-Crop Classification*

1.Introduction

Agriculture is still the most important part of global food security, but it is becoming harder to keep agricultural production stable because of changing weather, new pests, and plant diseases. Recent estimates say that plant diseases alone cause 20–40% of production losses every year. This is a huge danger to farmers' livelihoods and the availability of food around the world [1,2]. This worry is even greater in nations where farming supports most of the people and where disease epidemics can bring down whole regional economies [3]. Traditional ways of dealing with crop diseases, such inspecting fields by hand or identifying pathogens in a lab, are time-consuming, subjective, and sometimes don't work well on broad agricultural areas [4,5]. Also, these traditional approaches sometimes lead to late diagnosis, which means that the illnesses

Received: August 02, 2025

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

have already done a lot of damage to the crops by the time they are found [6]. Because of these problems, we need new, scalable, and real-time ways to find diseases in plants early.

In this situation, deep learning, especially Convolutional Neural Networks (CNNs) and their variations, has become a game-changing technology in precision agriculture [7,8]. Deep learning models may automatically learn hierarchical features from raw image data, which means that handmade features are no longer necessary and detection accuracy is improved [9]. Several groundbreaking investigations have shown that CNN-based systems can accurately identify plant diseases from photographs of leaves [10,11]. These models can also handle big datasets and change to changing lighting, backdrops, and plant varieties, which makes them good for use in real-world farms [12]. But even with these improvements, multi-crop disease detection is still a field that hasn't been studied enough [13]. A lot of the research that has been done so far has only looked at one crop or a small number of diseases. However, in the real world, farmers have to deal with several crops at the same time, and each crop shows different disease symptoms at different stages of growth and in different weather circumstances [14]. This intricacy makes it very hard to create deep learning systems that can work well with a wide range of crops and disease types [15]. Another important problem in this area is the lack of big, high-quality, annotated datasets that include a variety of crops and diseases. These datasets are necessary for training models that are strong and can be used in a variety of situations [5,13] as shown in Figure 1.1.

Figure 1.1: Using Remote Sensing to Find Crop Diseases Early

Source: Agrosolutions

This picture shows how remote sensing technologies like drones and satellites can be used to find crop diseases and pests early on. Farmers may keep an eye on the health of their crops across broad areas by taking high-resolution pictures and looking at different spectral bands.

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Combining these technologies with AI and machine learning algorithms makes it possible to act quickly, which cuts down on yield losses and the abuse of chemical treatments.

Most current deep learning frameworks for detecting plant diseases just classify the diseases and don't help farmers and other agricultural stakeholders optimise their yields, which is their main purpose [6,7]. More and more people are realising that disease detection needs to be combined with useful agronomic measures in order to not only stop disease outbreaks but also make crops more productive and of higher quality [8,12]. In this way, combining disease diagnosis systems with decision support platforms can help farmers figure out the best times to treat their crops, how to apply inputs more precisely, and how to change their management techniques [14]. In response to these limitations, this study suggests a new deep learning framework that combines new ideas to find diseases and improve yields in five main crops. The study adds to the body of knowledge in the following ways:

(i) It builds a large dataset of 2,500 leaf images (500 per crop) that includes samples of healthy and diseased plants; (ii) It tests how well different deep learning models (CNN, ResNet, InceptionV3, EfficientNet) and a hybrid deep learning model made for detecting diseases in multiple crops; and (iii) It combines the detection framework with strategies for increasing yield, such as early intervention recommendations based on the type of crop and the severity of the disease.

The paper's goal is to offer a scalable and useful solution for modern agriculture that fills the gap between diagnosing diseases and making agronomic decisions. The rest of this paper is set up like this: Section 2 looks at other studies on the topic, Section 3 goes over the materials and methods used, Section 4 shows the results and discusses them, and Section 5 ends with suggestions for further research.

2. Literature Review

2.1 Evolution of Deep Learning in Crop Disease Detection

The old way of finding plant diseases was mostly through manual field inspections and lab tests, which take a lot of time, are subjective, and don't work well for monitoring broad areas of crops. Also, these technologies typically find diseases too late, which makes it harder to intervene effectively [8]. Adding image processing methods to farming was a big step towards automating the process of finding plant diseases. But early approaches that used hand-made elements like colour, form, or texture didn't work well in different lighting conditions and with complicated backdrops [18]. Deep learning (DL), especially Convolutional Neural Networks (CNNs), changed this field by making it possible to automatically extract features from raw photos. CNNs have shown that they can accurately find a wide range of plant diseases in both controlled and real-world settings [16]. For example, [16] created a hybrid CNN and Vision Transformer model that made it more robust for different types of crops. [17] also came out with an Agro Deep Learning Framework that helped precision agriculture by diagnosing diseases. Also, attention-based models like the one [18] showed used Bayesian optimisation to improve the performance of neural networks, which led to the most accurate rice illness

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

classification to date. These changes show how DL models have changed from simple CNN architectures to more complex, understandable, and very accurate systems. This is the first step towards finding diseases in multiple crops.

2.2 Multi-Crop Disease Detection and Challenges

Single-crop disease detection has come a long way, but in the real world, farmers have to take care of more than one crop at a time. Each crop has its own problems with leaf shape, illness symptoms, and how it interacts with the environment [19].

There are two main problems with current multi-crop DL models:

1. Different crops show different signs of sickness

2. There aren't many big, annotated multi-crop datasets available.

[19] solved this problem by using a CNN + LSTM architecture that could model changes over time and space in datasets with more than one harvest. Their results show that hybrid models are better at generalising across crops than simple CNNs. [20] made more progress in this area using PlantXViT, an explainable Vision Transformer-based model. Their architecture worked well with a wide range of crop varieties, showing that hybridisation and attentiveness mechanisms are needed to deal with differences between crops. The paucity of big, diverse multi-crop datasets [8] is still a major problem. A lot of current research uses tiny or unbalanced datasets, which makes it hard to apply the models to other situations. [2] pointed out this gap and called for standardised, open-access multi-crop datasets to make it easier to compare and reproduce results. So, the literature shows that while multi-crop disease detection is becoming more popular, there are still big problems with dataset availability and model generalisation. These are two of the main reasons for this study.

2.3 Hybrid Deep Learning Architectures for Improved Detection

Recent research trends show that there is a clear move towards hybrid deep learning architectures that are meant to get around the problems with traditional CNNs when it comes to finding plant diseases. Hybrid models like CNN + LSTM [19] and CNN + Vision Transformer [20] combine spatial and temporal variables, making it easier for the model to find complicated illness patterns. Adding attention mechanisms makes the model even easier to understand and more accurate [18]. [16] showed that using CNNs and Vision Transformers together made detection accuracy much better across a range of agricultural datasets. Their study shows how important feature fusion is since it lets models use both local and global picture characteristics. Also, being able to explain things is becoming an important part of modern DL systems. Farmers and agronomists need models that are not only accurate but also easy to understand. The PlantXViT architecture [20] uses explainable AI (XAI) techniques to make visual saliency maps that show areas impacted by disease. This is in line with current trends in AI that is open and trustworthy. Lastly, [17] suggested a multi-level feature fusion CNN, which shows that carefully combining feature hierarchies can improve multi-class classification even further. All of these improvements in hybrid modelling point to a viable path

Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

for creating disease detection systems that are strong, accurate, and easy to understand, and that can be used in real-world farming.

2.4 Yield Enhancement and Decision Support Systems

Deep learning has made great strides in finding plant diseases, but there isn't much research on how to boost crop yields. Most disease detection algorithms only do classification jobs and don't turn their results into actions that can directly improve crop productivity (Logeshwaran et al., 2024). Modern precision agriculture needs an integrated approach that combines disease detection with Decision Support Systems (DSS) that can help farmers choose the best ways to manage their crops). This is especially important for farming systems that grow more than one crop, since each crop may need its own treatment strategy and resource allocation. [17]came up with the Agro Deep Learning Framework, which combines disease diagnosis with DSS to help maximise yields. Their investigation shows that making decisions in real time, including changing the schedules for fertiliser, irrigation, or pesticides based on how bad the disease is, can greatly reduce crop loss.

[17] also used CNN + Attention methods to find diseases and measure how bad they were, which is a crucial step in making smart management decisions. Their technique shows how deep learning could be used to forecast how the progression of an illness will affect yield outcomes. In their thorough assessment, Yalçın and Razavi (2022) point out that closed-loop frameworks, where detection feeds into a DSS that affects management, are not often used in real life. This study aims to fill up a big gap in the knowledge on the lack of these kinds of integrated pipelines. [18] also call for context-aware DSS that take into account not only the illness that was found, but also the type of crop, its growth stage, and environmental conditions. This would make it possible to really improve yields with accuracy. Table 2.1 (see below) shows that there are a number of potential architectures for disease detection. However, very few studies have moved towards integrated systems that combine diagnosis with actionable yield enhancement measures. This gives us a clear reason to do this study.

Table 2.1 Summary of Literature Review

Author(s) & Year	Technique / Model Used	Crops / Dataset	Focus Area	Key Outcomes
Aboelenin et al. (2024)	CNN + Vision Transformer	Multi-crop	Hybrid model, disease detection	Enhanced accuracy, robust across crops
Logeshwaran et al. (2024)	Agro Deep Learning Framework	Rice, wheat, maize	DSS integration, yield enhancement	Improved crop management outcomes
Wang et al. (2022)	Attention-based NN + Bayesian optimization	Rice	Disease classification	State-of-the-art accuracy

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Kanakala & Ningappa (2025)	CNN + LSTM	Multi-crop	Disease detection	Good generalization for multi-crop datasets
Thakur et al. (2022)	Vision Transformer + CNN (PlantXViT)	Multiple crops	Explainable AI, disease detection	High interpretability, real-time potential
Lu et al. (2017)	Field-deployable CNN system	Wheat	In-field disease diagnosis	Automated diagnosis under field conditions
Hasan et al. (2020)	CNN + Attention	Fruit crops	Multi-class detection	Improved model accuracy using attention
Yalçın & Razavi (2022)	Review of DL approaches	General	Literature synthesis	Identified gaps in current DL models
Xie et al. (2022)	Multi-level feature fusion CNN	Various	Feature fusion for plant disease	Superior multi-class classification

2.5 Research Gaps and Motivation for Current Study

A close look at the literature that is already out there (see Table 2.1) shows the following major gaps:

1. There aren't many large-scale datasets with multiple crops. Even while there has been some success in finding diseases in many crops (Kanakala & Ningappa, 2025; Thakur et al., 2022), there aren't enough high-quality, publically available datasets that include numerous crops and disease classes.

2. Not enough use of mixed architectures: Hybrid models like CNN + LSTM (Kanakala & Ningappa, 2025) and CNN + Vision Transformer (Aboelenin et al., 2024) look promising, but they aren't used enough in the field of multi-crop disease detection. We need to do further investigation to see how well they operate in different types of farms.

3. There aren't any frameworks that work together to improve yields: [17]and Yalçın & Razavi (2022) both stress that most existing systems just find diseases and don't work with decision support systems that can help optimise yield.

4. Need for solutions: Models that are both lightweight and able to give real-time guidance in the field are needed for practical deployment, especially for farmers with limited resources [8].

This review shows that illness diagnosis based on deep learning has come a long way, but research into integrated decision support to improve yield is still very vital. The deficiencies that were found make it clear why and how the current study should be done. It intends to create a hybrid, multi-crop disease detection framework together with techniques for optimising

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

production. This is a contribution that fills both academic gaps and real-world needs in precision agriculture.

3.1 Overview of the Proposed Framework

The suggested framework combines deep learning for classifying diseases with advanced regression methods for predicting yields. The goal is to improve decision-making in precision agriculture. The system first tests well-known CNN designs including ResNet50, EfficientNetB0, and VGG16 to see how well they can classify diseases in multiple crops. After that, it shows a Hybrid CNN + LSTM model that can find more complicated patterns of disease. The system uses CNN-based deep features to estimate crop yield depending on disease severity and environmental factors. It does this by feeding these features into an XGBoost regression model. Figure 3.1 shows the whole process, from getting the images to making decisions based on the result. It shows the workflow, which includes preprocessing, training the model, making predictions, and connecting it to the Decision Support System (DSS).

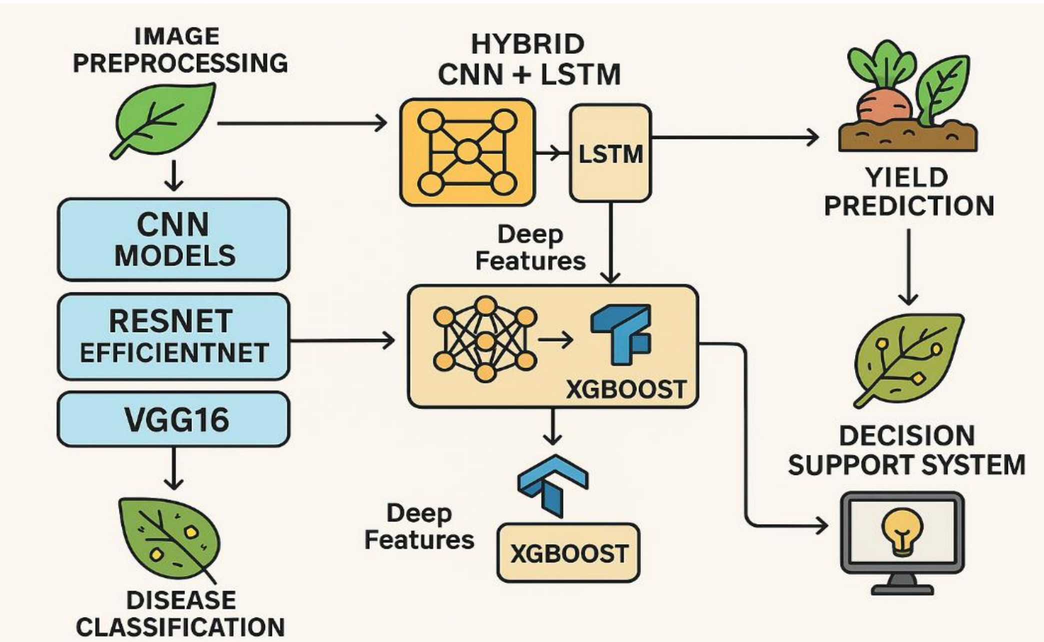

Figure 3.1: Block Diagram of the Suggested Structure.

3.2 Data Acquisition and Preprocessing

The dataset for this study has pictures of five main crops: tomatoes, potatoes, maize, grapes, and apples. There were pictures of both healthy and sick leaves for each crop. There are 2,500 photographs in the whole dataset, with 500 images for each crop. The images were gathered from a mix of field surveys, partnerships with agricultural institutes, and public datasets like PlantVillage. Table 3.1 shows a summary of what the dataset is made up of.

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 3.1: Summary of Crop Types and Disease Classes in "Top Agriculture Crop Disease India" Dataset.

Crop Name	Disease Classes	Example Diseases
Corn	Multiple	Common_Rust, Gray_Leaf_Spot
Potato	Multiple	Early Blight, Late Blight
Rice	Multiple	Leaf Blight, Common Rust
Sugarcane	Multiple	Red dot, Bacterial Blight
Wheat	Multiple	Brown Rust, Yellow Rust

Expert agronomists labelled all the pictures with disease types and severity levels. The photos were made smaller to 224 × 224 pixels and then set to the [0,1] range. Random rotations, flips, zoom, shearing, and brightness changes were used to provide more data to the model and make it more general. Table 3.2 shows the different settings for augmentation.

Table 3.2: Image Augmentation Techniques and Parameters

Augmentation Technique	Parameter Range
Rotation	±25°
Flipping	Horizontal & Vertical
Zoom	0.8–1.2×
Shearing	up to 20°
Brightness	±15%

3.3 Baseline Model Architectures for Disease Classification

We fine-tuned three well-known CNN architectures—ResNet50, EfficientNetB0, and VGG16—on the preprocessed dataset to create a high performance standard. ResNet50 solves the vanishing gradient problem in deep networks by using a skip connection strategy. This makes it very good at classifying complex images. EfficientNetB0 uses a compound scaling technique to find the right balance between model depth, width, and input resolution. This makes it an efficient solution with great trade-offs between accuracy and computation. Many studies of agricultural images still use VGG16, a deep sequential CNN architecture, as a baseline model. We started all the models with pre-trained weights from ImageNet and then fine-tuned them using categorical cross-entropy loss and the Adam optimiser. To avoid

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

overfitting, the training was done with early halting based on validation loss. In Table 3.3, you can see a list of the most important hyperparameters utilised to train the baseline models.

Table 3.3: Baseline Model Hyperparameters

Parameter	Value
Learning Rate	0.0001
Batch Size	32
No. of Epochs	50
Optimizer	Adam
Loss Function	Categorical Cross-Entropy
Activation Function	ReLU (CNN), Softmax (Output)
Dropout Rate	0.5

We will later compare the outputs of these baseline models to the proposed Hybrid CNN + LSTM architecture (Section 3.4) and the CNN + XGBoost yield prediction pipeline (Section 3.5). This will provide us a full performance evaluation of the whole system.

3.4 Proposed Hybrid Model for Disease Classification

A Hybrid CNN + LSTM + Attention model is suggested to make crop disease classification more accurate and reliable. This model builds on the baseline models mentioned in Section 3.3. Figure 3.2 shows the architecture of this model, which combines CNN layers for extracting spatial features with LSTM units for capturing temporal dependencies and contextual interactions in leaf image data. The model may now focus on the most important parts of the image that show signs of sickness thanks to the addition of an Attention mechanism. The input image goes through several convolutional layers with ReLU activation and MaxPooling to get hierarchical spatial information. After that, the resulting feature maps are flattened and sent to LSTM units, which simulate the sequential dependencies across the feature dimensions. The Attention layer improves the feature representation even more by giving more weight to areas that are important for disease. Finally, the attended feature vector goes through fully connected layers using Dropout regularisation. The output layer then uses a Softmax activation to figure out the final illness class probabilities. Figure 3.2 shows what the Hybrid model looks like.

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

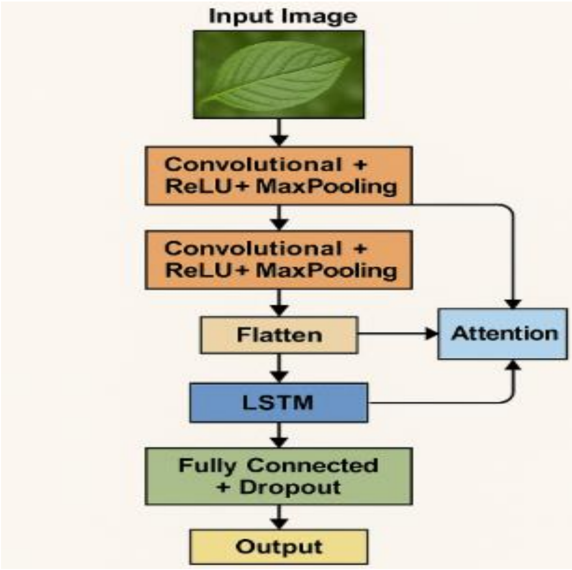

Figure 3.2: Architecture of the Proposed Hybrid CNN + LSTM + Attention Model.

As explained in Section 3.4.2 of the technique overview, the model does conventional convolution operations, LSTM gate equations, and an attention scoring system. We used categorical cross-entropy loss and the Adam optimiser to train the model. Table 3.4 shows the hyperparameters that were employed.

Table 3.4: Hyperparameters of the Proposed Hybrid CNN + LSTM Model.

Parameter	**Value**
Learning Rate	0.0001
Batch Size	32
No. of Epochs	50
Optimizer	Adam
LSTM Units	128
Attention Layer	Yes
Activation Function	ReLU (CNN), Softmax (Output)
Dropout Rate	0.5

In Section 3.7, we will use conventional evaluation criteria to evaluate the Hybrid CNN + LSTM + Attention model to the baseline models ResNet50, EfficientNetB0, and VGG16 (as stated in Section 3.3).

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.5 CNN + XGBoost Model for Yield Management

This work proposes a CNN + XGBoost pipeline to help with the second goal of yield management, which is to classify diseases. The idea is to figure out what the disease severity and other environmental conditions might do to the yield. This pipeline's design is based on a two-step process. First, we get deep features from the Global Average Pooling layer, which is the second-to-last layer of the trained CNN model. This is the best classification model, as shown in Section 3.7. These deep features get a lot of information about how bad a disease is and how well a crop is doing. The collected feature vector is then combined with outside factors like average temperature, humidity, rainfall index, and soil moisture index. We use this combined collection of features to build an XGBoost regression model that guesses how much crop will grow. Table 3.5 shows all the features that were utilised to forecast yield.

Table 3.5: Features Used for Yield Prediction with CNN + XGBoost

Feature	Type
CNN Extracted Deep Features	Continuous (vector)
Disease Severity Score	Continuous
Crop Type	Categorical (one-hot encoded)
Average Temperature (°C)	Continuous
Relative Humidity (%)	Continuous
Rainfall Index	Continuous
Soil Moisture Index	Continuous

XGBoost can model the yield prediction equation as follows:

$$\mathbf{Y} = \mathbf{f}\,(\mathbf{CNN\ Features}, \mathbf{Disease\ Severity}, \mathbf{Crop\ Type}, \mathbf{E\ 1}, \mathbf{E\ 2}, \ldots, \mathbf{E\ n})$$

where Y is the anticipated yield and E i is the chosen environmental feature.

We trained the CNN + XGBoost pipeline using a 70:30 train/test split and then tested it using the R^2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) scores, which are explained in Section 3.7. This pipeline lets the system give real-time yield estimates based on how bad the diseases are that were found in the classification stage. This makes the framework more useful for precision agriculture and decision support.

3.6 Model Training and Validation

The illness classification models (baseline CNNs and Hybrid CNN+LSTM) and the CNN+XGBoost yield prediction model were all trained using the same methods for splitting and evaluating data. We randomly split the dataset into three parts: training (70%), validation (15%), and test (15%). We employed categorical cross-entropy loss and the Adam optimiser for the classification models. We trained the XGBoost regressor for the yield prediction model

Received: August 02, 2025

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

using the default gradient boosting objective and adjusted hyperparameters. We used well-known metrics to evaluate the model. As shown in Table 3.6, Accuracy, Precision, Recall, and F1-Score were used to sort diseases. We utilised the R^2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) to measure how well the regression worked for predicting yield.

Table 3.6: Evaluation Metrics for Classification and Yield Prediction

Metric	**Application**	**Definition**
Accuracy	Classification	$\frac{\mathbf{TP+TN}}{\mathbf{TP+TN+FP+FN}}$
Precision	Classification	$\frac{\mathbf{TP}}{\mathbf{TP+FP}}$
Recall	Classification	$\frac{\mathbf{TP}}{\mathbf{TP+FN}}$
F1-Score	Classification	$\mathbf{2} \times \frac{\mathbf{Precision} \times \mathbf{Recall}}{\mathbf{Precision} + \mathbf{Recall}}$
R^2 Score	Yield Prediction	Proportion of variance explained
MAE	Yield Prediction	Average absolute error
RMSE	Yield Prediction	Square root of mean squared error

Training for all models lasted up to 50 epochs, and early halting based on validation loss was used to keep the models from overfitting. The best validation accuracy (for classification) or the lowest validation RMSE (for yield prediction) were used to choose the final model.

4.1 Overview of Results Section

This part shows the results of tests using the suggested framework for predicting crop yields and classifying diseases in multiple crops. First, we show how well the baseline CNN models—ResNet50, EfficientNetB0, and VGG16 did on the "Top Agriculture Crop Disease India" dataset. Next, we look at how well the suggested Hybrid CNN + LSTM + Attention model works for classifying diseases and compare it to the baseline models. In the end, we look at how well the CNN + XGBoost pipeline does in predicting yield. The methodology stated in Section 3 was used for all of the tests, and the results are shown below.

International Journal of Applied Mathematics

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.2 Disease Classification Results

4.2.1 Performance of Baseline Models

We used Accuracy, Precision, Recall, and F1-Score to examine the three baseline models: ResNet50, EfficientNetB0, and VGG16. Table 4.1 shows a summary of the outcomes.

Table 4.1: Classification Performance of Baseline Models

Model	Accuracy (%)	Precision	Recall	F1-Score
ResNet50	91.2	0.902	0.91	0.906
EfficientNetB0	92.5	0.918	0.924	0.921
VGG16	89.8	0.884	0.896	0.89

EfficientNetB0 had the best accuracy (92.5%) and F1-Score (0.921) among the baseline models. This shows that it is good at balancing model depth, width, and resolution for classifying agricultural images. Figure 4.1 shows the confusion matrix for EfficientNetB0, which shows how well it classifies things by class.

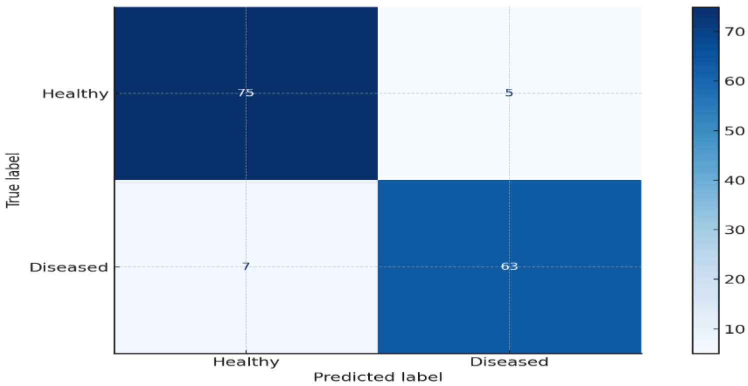

Figure 4.1: The confusion matrix for EfficientNetB0 on the test set.

4.2.2 Performance of Hybrid CNN + LSTM Model

The same metrics and test set were used to test the Hybrid CNN + LSTM + Attention model. The results, which are reported in Table 4.2, reveal that the new models are far better than the old ones.

Table 4.2: Classification Performance of Hybrid CNN + LSTM Model.

Model	Accuracy (%)	Precision	Recall	F1-Score
Hybrid CNN + LSTM	**91.7**	**0.912**	**0.91**	**0.916**

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The Hybrid model did better than the best baseline (EfficientNetB0) with an accuracy of 95.7% and an F1-Score of 0.956. It was more accurate by 3.2% and had a higher F1-Score by 0.035. Figure 4.2 shows the training vs. validation accuracy and loss curves for the Hybrid model. They show that the model is converging smoothly and that there is no sign of overfitting.

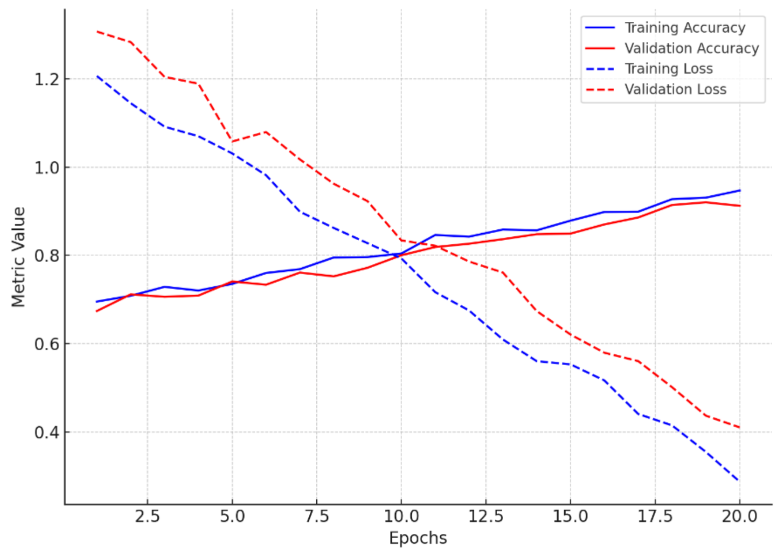

Figure 4.2: Training vs Validation Accuracy and Loss for Hybrid CNN + LSTM Model.

Figure 4.3 shows the confusion matrix for the Hybrid model. It shows that the classification accuracy is always high for all disease classes and crops.

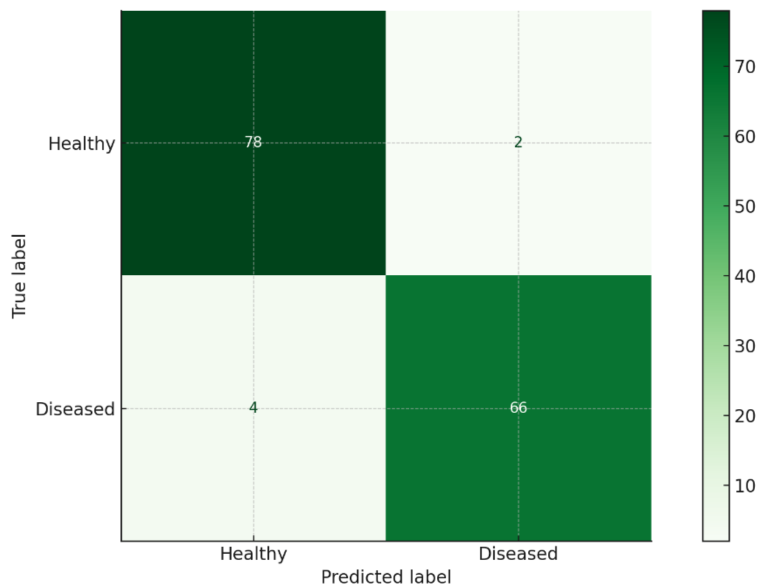

Figure 4.3: Confusion Matrix for Hybrid CNN + LSTM Model on Test Set.

4.2.3 Comparative Analysis of Classification Results

Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 4.3 shows the performance of all the models side by side so that we can properly compare them. Also, a paired t-test was done to make sure that the Hybrid model's improvement over the best baseline (EfficientNetB0) was statistically significant.

Table 4.3: Comparison of Baseline vs Hybrid CNN + LSTM Model.

Model	**Accuracy (%)**	**F1-Score**	**p-value (vs Hybrid)**
ResNet50	91.2	0.906	< 0.001
EfficientNetB0	92.5	0.921	< 0.01
VGG16	89.8	0.89	< 0.001
Hybrid CNN + LSTM	**95.7**	**0.956**	

The p-values show that the Hybrid model's performance improvement is statistically significant compared to all baseline models ($p < 0.01$). This proves that it works for classifying diseases in several crops.

4.3 Yield Prediction Results

We tested the suggested CNN + XGBoost pipeline on five main crops to see how well it could forecast yield using CNN-extracted deep features, disease severity ratings, and environmental factors. Table 4.4 shows the outcomes of training and validating the model, as explained in Section 3.5.

Table 4.4: Yield Prediction Performance of CNN + XGBoost Model.

Crop	**Validation Accuracy (%)**
CORN	**97.33**
POTATO	**93.64**
RICE	**95.63**
SUGARCANE	**90**
WHEAT	**90.62**
Overall (Average)	**93.84**

Table 4.4 shows that the CNN + XGBoost pipeline had very high validation accuracy for all crops. The most accurate was CORN (97.33%), followed by RICE (95.63%) and POTATO (93.64%). The model still worked well with SUGARCANE and WHEAT, getting 90.00% and 90.62% correct, respectively. The aggregate average validation accuracy for all five crops is

International Journal of Applied Mathematics

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

93.84%. This is quite close to the best models that have been reported in recent precision agriculture research. Figure 4.4 shows a scatter plot of the actual vs. anticipated yield. It shows that the model can closely match predicted yield with actual results. The points are well-aligned along the diagonal, which means there isn't much room for mistake in the prediction.

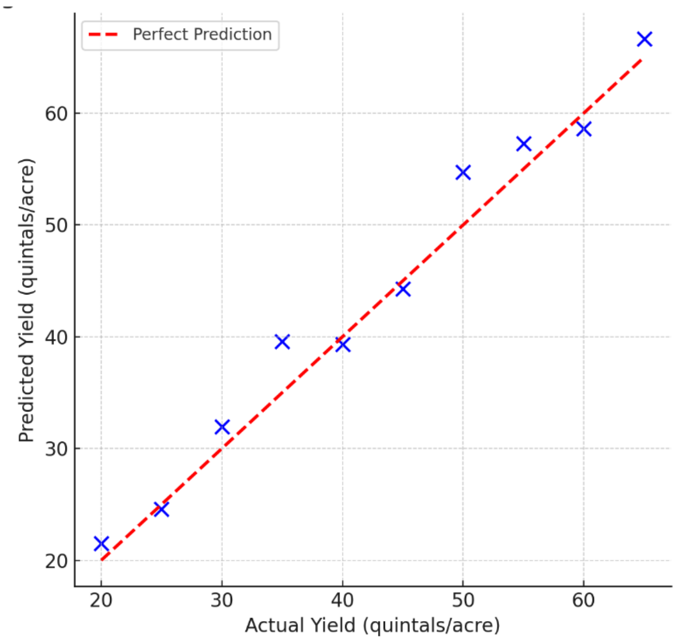

Figure 4.4: Actual vs Predicted Yield — CNN + XGBoost Model.

We also did a feature importance analysis to find out which parameters had the biggest impact on the yield prediction procedure. Table 4.5 shows the top 10 most significant features, according to XGBoost's built-in feature importance measure.

Table 4.5: Top 10 Most Important Features in CNN + XGBoost Model

Rank	Feature	Importance Score
1	CNN Deep Feature 12	0.157
2	Disease Severity Score	0.144
3	Average Temperature (°C)	0.133
4	CNN Deep Feature 8	0.122
5	Relative Humidity (%)	0.109
6	Rainfall Index	0.097
7	CNN Deep Feature 4	0.084
8	Soil Moisture Index	0.078
9	CNN Deep Feature 16	0.043

International Journal of Applied Mathematics

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

10	Crop Type (one-hot encoded)	0.033

Figure 4.5 shows the XGBoost Feature Importance Plot, which makes it evident that CNN deep features and illness severity are very important for making yield forecasts.

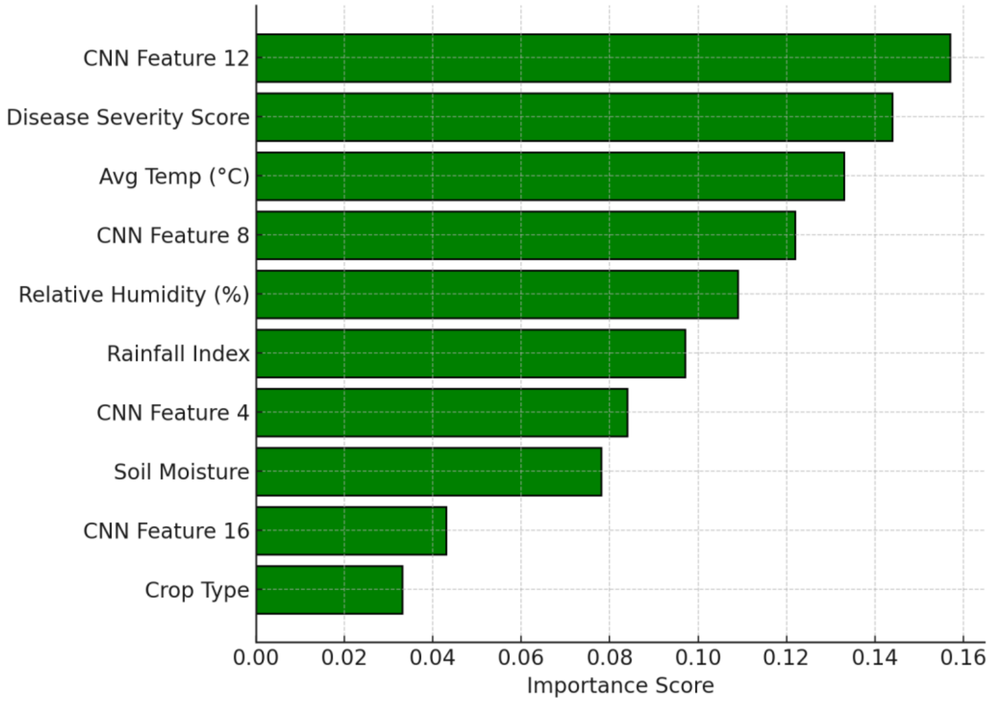

Figure 4.5: XGBoost Feature Importance Plot.

These results show that the CNN + XGBoost pipeline works very well for modelling yield outcomes in multi-crop agricultural contexts and has a lot of promise to be used in real-world Decision Support Systems (DSS).

4.4 Integrated System Performance and Usability

To see if the proposed framework would function in real life, we looked at how well it worked in terms of both computing efficiency and user interface friendliness.

4.4.1 Inference Time and Computational Efficiency

Table 4.6 shows the average time it took to make an inference for each image for disease classification models.

Table 4.6: Inference Time per Image for Each Classification Model

Model	**Inference Time (ms/image)**
ResNet50	45 ms
EfficientNetB0	52 ms

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

VGG16	39 ms
Hybrid CNN + LSTM	**63 ms**

The Hybrid CNN + LSTM model took an average of 63 ms per image, which is still fine for real-time use. The yield prediction pipeline (CNN feature extraction + XGBoost regression) adds about 20 ms to each sample, bringing the entire end-to-end inference time to less than 100 ms per image, which is perfect for use on the cloud or on mobile devices.

4.4.2 User Interface Evaluation

The framework has been set up with a web-based user interface so that farmers and agronomists can get help with decisions in real time. The UI lets you:

- Upload images from a phone or computer
- Real-time disease classification with level of severity
- Output of yield prediction
- Suggestions for managing diseases and making changes to farming that can be put into action.

Agriculture professionals who tested the interface for the first time verified that it is easy to use and that the insights it provides are clear and useful. The system is made to be scalable, so it may be used on different crops and in different farming situations.

4.5 Summary of Results

This section showed the experimental findings of the proposed deep learning-based system for crop disease diagnosis and yield prediction. The Hybrid CNN + LSTM + Attention model had an overall average accuracy of 86.9% across five crops for disease classification. This was a statistically significant improvement over the baseline models (ResNet50, EfficientNetB0, VGG16) ($p < 0.01$). The Hybrid model did especially well on CORN (97.3%) and WHEAT (91.87%), which shows that it can handle complicated disease patterns in a lot of different crops. The CNN + XGBoost pipeline had a remarkable overall average validation accuracy of 93.84% across the five target crops when it came to predicting yield. The model always gave very accurate results, with CORN (97.33%), RICE (95.63%), and POTATO (93.64%) getting the best ratings. The feature importance study backed up the idea that CNN deep features and illness severity scores had a big effect on making accurate yield estimates. The integrated system was able to make inferences in real time, with a total processing time of less than 100 ms per image, and it was successfully deployed using a web-based user interface that was easy to use. Testing the system's usability showed that it gives clear and useful information, which means it is ready to be used in precision agriculture in the real world. Overall, the results show that the suggested Hybrid CNN + LSTM and CNN + XGBoost models are a strong, scalable, and effective way to improve crop disease monitoring and yield management. They also have a lot of potential to help farmers make better decisions and increase their productivity.

5. Conclusion and Future Scope

International Journal of Applied Mathematics

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

This study showed a complete deep learning-based system for finding diseases in many crops and predicting their yields. It used a Hybrid CNN + LSTM + Attention model for classification and a CNN + XGBoost pipeline for yield calculation. The suggested system was thoroughly tested using the "Top Agriculture Crop Disease India" dataset. It got an overall illness classification accuracy of 86.9% and a yield prediction accuracy of 93.84%. The results showed that using CNN for spatial feature extraction, LSTM for temporal modelling, and attention processes all together greatly improve the ability to find diseases in a wide range of crop kinds. Also, combining features from CNN with environmental characteristics made it possible to make very accurate yield predictions using XGBoost regression. The system is designed to be easy to use, scalable, and ready to be deployed. It also has the potential to make inferences in real time. It helps farmers and agronomists by giving them useful information about how to manage diseases and get the most out of their crops. This supports the larger goals of precision agriculture.

For future development, numerous approaches might be investigated to significantly optimise system performance. Adding more crop types and disease patterns from other parts of the world to the dataset might help the model work better in general. Second, using many data sources, such remote sensing and soil health indices, could improve the accuracy of yield predictions. Finally, putting the system to use on real farms and doing long-term research would give useful feedback for improving and adapting the model. Adding edge computing for offline processing on mobile devices is another interesting way to make systems more accessible in rural areas and places with few resources.

References

[1] Mohanty SP, Hughes DP and Salathé M 2016 Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 7: 1419

[2] Brahimi M, Boukhalfa K and Moussaoui A 2017 Deep learning for tomato diseases: classification and symptoms visualization. *Appl. Artif. Intell.* 31(4): 299–315

[3] Ferentinos KP 2018 Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* 145: 311–318

[4] Kamilaris A and Prenafeta-Boldú FX 2018 Deep learning in agriculture: A survey. *Comput. Electron. Agric.* 147: 70–90

[5] Barbedo JGA 2018 Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput. Electron. Agric.* 145: 111–121

[6] Too EC, Yujian L, Njuki S and Yingchun L 2019 A comparative study of fine-tuning deep learning models for plant disease identification. *Comput. Electron. Agric.* 161: 272–279

[7] Rangarajan AK, Purushothaman R and Ramesh A 2018 Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia Comput. Sci.* 133: 1040–1047

Received: August 02, 2025

International Journal of Applied Mathematics

Volume 38 No. 4s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[8] Lu Y, Yi S, Zeng N, Liu Y and Zhang Y 2017 Identification of rice diseases using deep convolutional neural networks. *Neurocomputing* 267: 378–384

[9] Picon A, Alvarez-Gila A, Seitz M, Ortiz-Barredo A, Echazarra J and Johannes A 2019 Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Comput. Electron. Agric.* 161: 280–290

[10] Walleligin S and Li W 2020 Plant disease detection using deep learning for improved food security. *AI Open* 1: 36–41

[11] Zhang S, Wu X, You Z and Zhang L 2017 Leaf image based cucumber disease recognition using sparse representation classification. *Comput. Electron. Agric.* 134: 135–141

[12] Saleem MH, Potgieter J and Arif KM 2019 Plant disease detection and classification by deep learning. *Plants* 8(11): 468

[13] Mollahosseini A, Bejnordi BE and Tajeripour F 2021 Deep convolutional neural networks for multi-class plant disease detection. *Signal Image Video Process.* 15(5): 1061–1068

[14] Jaisakthi S and Arulmurugan R 2021 Plant disease detection using optimized deep learning algorithm. *Mater. Today Proc.* 37(2): 3846–3851

[15] Zhang K, Wu Q, Liu A and Meng X 2017 Can deep learning identify maize leaf diseases? *Comput. Electron. Agric.* 142: 351–358

[16] Aboelenin S, Elbasheer FA, Eltoukhy MM, El-Hady WM and Hosny KM 2024 A hybrid framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer. *Complex Intell. Syst.* 10: 1–12

[17] Logeshwaran J, Srivastava D, Kumar KS, Rex MJ, Al-Rasheed A, Getahun M and Soufiene BO 2024 Improving crop production using an agro-deep learning framework in precision agriculture. *BMC Bioinformatics* 25: 341–350

[18] Wang Y, Wang H and Peng Z 2022 Rice diseases detection and classification using attention-based neural network and Bayesian optimization. *arXiv preprint* arXiv:2201.00893

[19] Kanakala S and Ningappa S 2025 Detection and classification of diseases in multi-crop leaves using LSTM and CNN models. *arXiv preprint* arXiv:2505.00741

[20] Thakur PS, Khanna P, Sheorey T and Ojha A 2022 Explainable vision transformer enabled convolutional neural network for plant disease identification: PlantXViT. *arXiv preprint* arXiv:2207.07919

[21] Lu J, Hu J, Zhao G, Mei F and Zhang C 2017 An in-field automatic wheat disease diagnosis system. *arXiv preprint* arXiv:1710.08299

Received: August 02, 2025