Exploration of Foods and Foodomics

Open Exploration

Open Access Original Article

Unified CNN-based recognition of multi-crop leaf diseases with web deployment

Zarif Wasif Bhuiyan[1,2*], Zaed Bin Monir Atib[1,2], Mahady Hasan[3], Md. Tarek Habib[1,2]

[1]Department of Computer Science and Engineering, Independent University, Bangladesh (IUB), Dhaka 1245, Bangladesh
[2]Fab Lab IUB, Independent University, Bangladesh (IUB), Dhaka 1245, Bangladesh
[3]Department of Computer Science and Engineering, Asian University for Women, Chattogram 4000, Bangladesh

***Correspondence:** Zarif Wasif Bhuiyan, Department of Computer Science and Engineering, Independent University, Bangladesh (IUB), Dhaka 1245, Bangladesh. zarifwasif20@gmail.com
Academic Editor: Gavin Walker, University of Limerick, Ireland
Received: October 13, 2025 **Accepted:** September 7, 2026 **Published:** November 10, 2026

Cite this article: Bhuiyan ZW, Atib ZBM, Hasan M, Habib MT. Unified CNN-based recognition of multi-crop leaf diseases with web deployment. Explor Foods Foodomics. 2026;4:1010200. https://doi.org/10.37349/eff.2026.1010200

Abstract

Aim: Timely and accurate identification of crop leaf diseases is crucial for food security and sustainable agriculture. This study aims to develop a multi-crop leaf disease recognition framework that covers corn, potato, rice, and tomato, addressing the challenges associated with maintaining high performance under diverse agricultural imaging conditions.

Methods: We utilized a balanced dataset of 21,000 curated leaf images spanning fourteen classes, including healthy and diseased samples. A custom convolutional neural network (CNN) was specifically designed to capture fine-grained lesion patterns. The model was trained using a unified pipeline with standardized data splits and preprocessing, ensuring reproducible evaluation. The performance was compared against standard deep learning architectures (ResNet-50, VGG-16) and traditional machine learning classifiers, including *k*-nearest neighbors (*k*-NN), support vector machine (SVM), and Random Forest.

Results: The custom CNN achieved near-perfect classification accuracy on the test set, outperforming ResNet-50, VGG-16, and traditional machine learning classifiers under identical conditions. While the model showed potential for generalization to diverse field scenarios, the evaluation was limited to curated datasets. Additionally, the trained model was deployed in a web application to demonstrate real-time inference and practical utility for leaf disease assessment.

Conclusions: This study establishes a robust and reproducible benchmark for multi-crop disease classification. The results highlight the model's effectiveness in precision agriculture, supporting early intervention strategies, though further validation in field conditions is needed to confirm generalization.

Keywords

multi-crop leaf disease, custom CNN, deep learning, precision agriculture, plant disease recognition, web deployment

© The Author(s) 2026. This is an Open Access article licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Introduction

Detecting crop leaf diseases quickly and correctly is essential for food security worldwide, for the profitability of farms, and for the sustainability of agriculture. However, identifying diseases from field images remains extremely challenging due to widely varying lighting, backgrounds, and the often-similar appearance of different foliar symptoms. Manual field scouting, the conventional approach, is time-consuming and gives inconsistent results, so treatments are often applied late or incorrectly, which can reduce yields and cause excessive pesticide use. While deep learning models have achieved near-perfect accuracy on curated lab datasets, they frequently falter in uncontrolled field conditions, revealing a persistent mismatch between benchmark results and practical use. With smartphone cameras now present on almost every farm but reliable diagnostic models still scarce, this mismatch has become a major obstacle. If left unaddressed, farmers will keep depending on subjective visual checks, resulting in prolonged disease outbreaks, higher crop losses, and diminished confidence in digital agriculture tools. This research therefore aims to narrow this performance gap through a dependable framework for identifying leaf diseases across multiple crops that delivers consistent, high-accuracy diagnoses from single leaf images under diverse field conditions. The study focuses on four major crops, namely corn, potato, rice, and tomato, covering a broad range of healthy and diseased leaf classes. By focusing on a generalizable solution across multiple crops, we intend to facilitate earlier disease intervention and strengthen sustainable crop management. Our hypothesis is that a custom convolutional neural network (CNN), when trained on a balanced multi-crop dataset, will achieve higher classification accuracy and generalization than both standard deep architectures and traditional machine learning classifiers under controlled evaluation conditions.

To tackle these challenges, we propose a unified, AI-based approach to crop disease recognition that places equal emphasis on accuracy and robustness in the field. This study builds upon existing open-source datasets [1–3] by assembling 21,000 images from four important crops, namely corn, potato, rice, and tomato, into a balanced collection in which healthy and diseased classes are equally represented. Although the images themselves are not newly collected, what is new in this work is a standardized, unified pipeline that supports rigorous and reproducible assessment of both deep learning models and traditional classifiers across several crops. Building on this foundation, we designed a custom CNN whose architecture is adapted to capture fine-grained lesion details in leaf images [4, 5]. While the architecture (five 3 × 3 convolutional blocks with batch normalization and ReLU activations) follows standard CNN design principles, its domain-specific optimization for leaf lesion recognition enables superior classification performance across multiple crops and disease categories. Every model was trained and tested within the same pipeline, which included standardized preprocessing, augmentation, and training protocols to eliminate confounding factors and enable a rigorous comparison [4–6]. Under identical conditions, our custom CNN markedly outperformed leading deep transfer-learning models, namely ResNet-50 and VGG-16, as well as the traditional classifiers *k*-nearest neighbors (*k*-NN), support vector machine (SVM), and Random Forest. On the test set, the custom CNN reached over 99% classification accuracy, and its precision and recall stayed high for every crop and disease category. These results underscore the value of using balanced data and domain-specific model design to achieve reliable, reproducible performance, while avoiding overstatement about field deployment [7–10]. Furthermore, we deployed the trained model in a web-based application and validated its real-time prediction capability on new leaf images, thereby confirming the practical utility of our approach for on-site disease diagnostics [11–14]. In summary, this work establishes a new benchmark for multi-crop disease identification by delivering a high-accuracy diagnostic framework evaluated on curated datasets that can accelerate disease detection and support precision agriculture [15–17].

The following are the primary contributions of our work:

1. To create a 14-class multi-crop leaf image dataset encompassing diseased and healthy leaves of corn, potato, rice, and tomato, enabling robust model training and evaluation for disease classification.

2. To develop a custom CNN model specifically optimized for multi-crop leaf disease detection, achieving superior feature extraction and classification performance compared to standard CNN models.
3. To conduct a comprehensive comparative evaluation against traditional machine learning classifiers (*k*-NN, SVM, Random Forest), demonstrating the proposed CNN's superior accuracy and robustness in multi-class leaf disease classification.
4. To achieve consistently high classification accuracy across diverse disease categories and crops, validating the model's generalization capability and reliability for broad-spectrum plant disease detection.
5. To deploy the trained model in a real-time web application for practical use, providing an accessible tool for instant leaf disease diagnosis and decision support in agricultural settings.

Pan et al. [4] built a lightweight YOLOv8-based detector for rice leaf diseases that achieved 95% performance with a model size of only 6 MB. By coupling NASNetMobile features with few-shot learning in a hybrid framework, Saleem et al. [5] obtained 98% accuracy across several crops. Tunio et al. [6] fused CNN and Transformer representations and applied adversarial domain adaptation, which raised accuracy under field conditions by 13%. A compact, lesion-focused CNN inspired by the Inception design was introduced by Bouacida et al. [7], reaching 94% accuracy on PlantVillage and 97% on crops not seen during training. Islam et al. [8] reported that ResNet-50 achieved 99% accuracy on PlantVillage and integrated the model into a web application for real-time diagnosis. To enable edge deployment, Khan et al. [9] applied quantization to MobileNetV3-Small and retained 99.5% accuracy in real time on low-power devices. The EfficientNet-Transformer hybrid of Feng et al. [10] scored 99.8% on a curated dataset, yet its accuracy dropped to 86.9% on field images of cassava. Using multispectral leaf images, De Silva and Brown [11] found that the vision transformer (ViT) variant ViT-B16 delivered the best accuracy (83%). Hemalatha and Jayachandran [12] designed a multitask ViT that jointly classifies diseases and segments lesions, attaining 99.97% accuracy on PlantVillage together with near-perfect lesion masks. Baek [13] developed an ensemble of ViTs that takes multi-leaf inputs; the ensemble outperformed single models, reaching 99% accuracy on apple and grape and 96% on tomato. Rajagopal et al. [14] combined lesion segmentation and handcrafted features with a multi-kernel SVM classifier and obtained 96% accuracy on an apple leaf disease dataset. According to Ngugi et al. [15], fine-tuned CNNs can exceed 95% accuracy on laboratory datasets, whereas traditional methods break down when images contain complex variations. Demilie [16] also observed that modern CNNs and one-stage detectors perform almost perfectly in controlled settings, but their performance deteriorates considerably under real field conditions.

Nevertheless, several important gaps persist. High accuracies are frequently reported on narrow datasets, typically covering a single crop or images with plain backgrounds, and such results may overestimate performance in practice. Models tuned under laboratory conditions tend to perform poorly in the field, yet few studies tackle this domain shift. Moreover, some highly accurate models rely on heavy architectures or ensembles that are impractical to deploy, whereas lighter models usually handle fewer classes. Fair comparison and reproducibility are further hampered by inconsistent data splits and imbalanced classes [17]. To overcome these limitations, this study offers a curated and balanced multi-crop dataset, together with a unified pipeline in which deep learning and traditional methods are evaluated under the same conditions. With equal representation of every class, fixed data splits, and uniform preprocessing, the study serves as a reproducible benchmark, although it does not claim direct field validation. The proposed custom CNN reaches state-of-the-art accuracy while keeping complexity modest, and the evaluation establishes a reproducible point of reference. In summary, this approach lays the groundwork for potential field deployment, while explicitly acknowledging that independent field trials are left for future work.

Materials and methods

The primary aim of this work is to tackle two key issues in crop disease recognition, high accuracy and robust generalization, through the design of a new CNN architecture that is systematically compared against deep learning and traditional machine learning baselines. As illustrated in Figure 1, the end-to-end workflow uses standardized preprocessing, balanced data splits, and uniform training protocols so that all models are assessed fairly. Alongside the proposed CNN, the experiments include ResNet-50 and VGG-16 as deep learning baselines and *k*-NN, SVM, and Random Forest as traditional classifiers. This broad comparison highlights the trade-offs between modern deep architectures and conventional machine learning techniques in multi-crop disease classification.

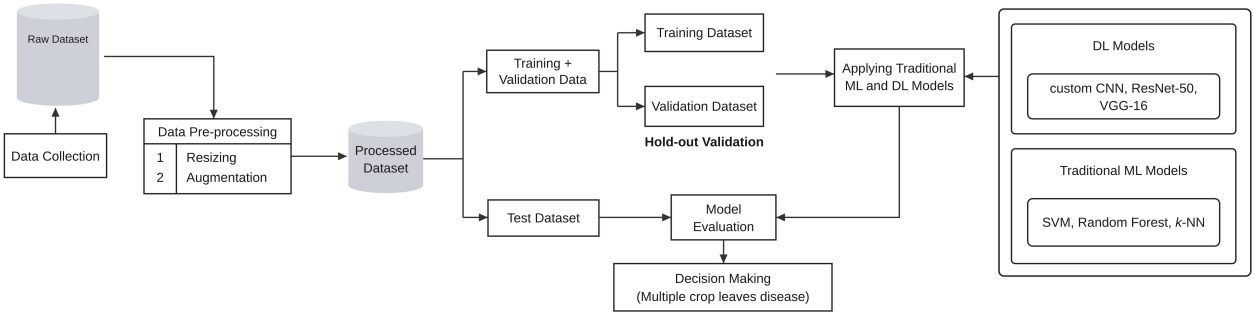


Figure 1. Overview of the proposed crop disease recognition pipeline. *k*-NN: *k*-nearest neighbors; SVM: support vector machine.

Dataset

We curated a balanced set of 21,000 high-resolution leaf images covering four major crops, namely corn, potato, rice, and tomato. Healthy and diseased samples are both included, yielding 14 classes in total. The diseases comprise leaf spot, common rust, and northern leaf blight in corn; early and late blight in potato; bacterial blight, brown spot, and leaf smut in rice; and bacterial spot and spider mite damage in tomato [1–3].

Because healthy and diseased samples are equally represented for every crop, this design reduces bias and enables reliable macro-level analysis. Representative examples are shown in Figure 2 and Table 1. Fixed train, validation, and test splits were used to ensure reproducibility and prevent data leakage. Cross-validation was not applied to maintain consistent evaluation across all models.

Data processing and model training workflow

All experiments were conducted using Python 3.10 with TensorFlow 2.15 and Keras 2.15. The models were trained on an NVIDIA GeForce RTX 3060 GPU (12 GB VRAM) using CUDA 12.2 and cuDNN 8.9 acceleration. Image preprocessing and augmentation were implemented using OpenCV 4.8 and TensorFlow image processing utilities. Data augmentation was applied uniformly across all models, including random rotations of up to ± 15°, horizontal and vertical flips, and brightness adjustments of ± 10% to improve model generalization. All models were trained using the same data splits, preprocessing procedures, and evaluation protocols to ensure a fair and reproducible comparison. Model training and evaluation were performed on a workstation equipped with an Intel Core i7 processor and 32 GB RAM.

We designed a foliar-specific CNN to capture multi-scale lesion cues under diverse conditions. The proposed backbone consists of five convolutional blocks, where the input size is 224 × 224 × 3. The number of filters in each block is as follows: Block 1: 32 filters, Block 2: 64 filters, Block 3: 128 filters, Block 4: 256 filters, and Block 5: 512 filters. The flattened feature maps are then connected to a dense layer followed by a 14-way Softmax output. The total number of trainable parameters in the model is approximately 2.3 million. Regularization techniques include dropout (0.25), label smoothing, and weight decay, while optimization uses categorical cross-entropy with a staged learning-rate schedule and early stopping. With

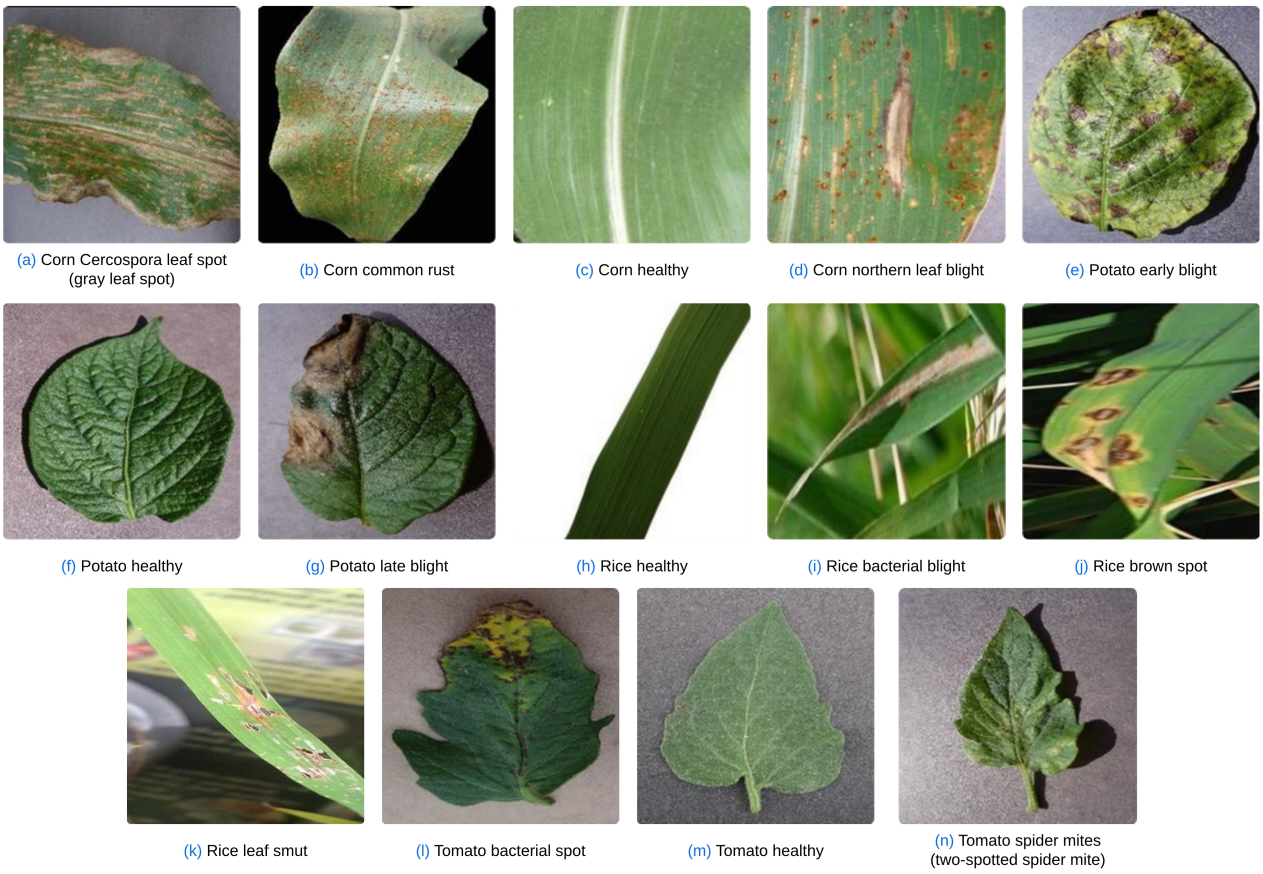


Figure 2. Sample images (a–n) illustrating healthy and diseased leaves for the four crops.

Table 1. Summary of the collected dataset.

Class name	Number of images			
	Training image usage	Validation image usage	Test image usage	Total image
Corn *Cercospora* leaf spot (gray leaf spot)	1,080	270	150	1,500
Corn common rust	1,080	270	150	1,500
Corn northern leaf blight	1,080	270	150	1,500
Corn healthy	1,080	270	150	1,500
Potato early blight	1,080	270	150	1,500
Potato late blight	1,080	270	150	1,500
Potato healthy	1,080	270	150	1,500
Rice bacterial blight	1,080	270	150	1,500
Rice brown spot	1,080	270	150	1,500
Rice leaf smut	1,080	270	150	1,500
Rice healthy	1,080	270	150	1,500
Tomato bacterial spot	1,080	270	150	1,500
Tomato spider mites (two-spotted)	1,080	270	150	1,500
Tomato healthy	1,080	270	150	1,500
Total images	15,120	3,780	2,100	21,000

input (W_1, H_1) and convolution parameters (F, P, S), where F is the filter size, P is the padding, and S is the stride, the output width W_2 and height H_2 are given by:

$$W_2 = \frac{W_1 - F + 2 \times P}{S} + 1 \quad (1)$$

$$H_2 = \frac{H_1 - F + 2 \times P}{S} + 1 \quad (2)$$

and pooling with window size F_s and stride S_s reduces the spatial size as:

$$W_2 = \frac{W_1 - F_s}{S_s} + 1 \tag{3}$$

Feature maps are flattened and passed to a dense layer, then a 14-way Softmax:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{14} e^{z_j}} \tag{4}$$

for class probabilities, where z_i is the logit of class i, z_j is the logit of class j (j = 1, ..., 14), and p_i is the predicted probability of class i. Regularization includes dropout (e.g., 0.25), label smoothing, and weight decay; optimization uses categorical cross-entropy with a staged learning-rate schedule and early stopping. The architecture's progressive depth and controlled down sampling stabilize small-lesion detection while preserving context, which—within our unified pipeline—drives its superior precision/recall across all crops-disease categories.

ResNet-50 employs residual learning to stabilize very deep optimization. Instead of directly approximating the target mapping $H(x)$, each residual block learns a residual function $F(x)$ and adds it to the input via an identity shortcut,

$$F(x) = H(x) - x \tag{5}$$

which yields through the skip connection:

$$H(x) = F(x) + x \tag{6}$$

These skip connections preserve gradient flow and mitigate vanishing/exploding gradients in 50-layer stacks. Our configuration ingests RGB inputs of 224 × 224 × 3, retains the standard stem and fine-tunes all convolutional stages from ImageNet initialization.

VGG-16 stacks small 3 × 3 convolutions to build depth efficiently, interleaved with max-pooling to down sample. Given input size (W_{in}, H_{in}), a convolution with kernel F, padding P, and stride S yields the output size (W_{out}, H_{out}):

$$W_{out} = \frac{W_{in} - F + 2P}{S} + 1 \tag{7}$$

$$H_{out} = \frac{H_{in} - F + 2P}{S} + 1 \tag{8}$$

followed by max-pooling that halves spatial dimensions. We fine-tune an ImageNet-pretrained VGG-16 on 224 × 224 × 3 leaf images, replacing the way head with a 14-class Softmax layer and retaining dropout in the fully connected blocks. To curb overfitting on leaf imagery, we use L2 regularization, moderate data augmentation, class-balanced sampling, and early stopping. This design captures fine textures (speckles, rust pustules) in early blocks and progressively aggregates larger lesion structures in deeper layers, yielding a strong, reproducible baseline under our unified training protocol.

SVM separates classes by maximizing the margin in a (possibly) high-dimensional feature space via kernels. Specifically, the features consist of raw pixel values from resized leaf images of 224 × 224. These pixel vectors are then reduced using principal component analysis (PCA) to the top 20 components and standardized using StandardScaler. This PCA-reduced, standardized representation is used as input for SVM, k-NN, and Random Forest classifiers to ensure reproducible and consistent evaluation. We tune C, kernel [linear vs. radial basis function (RBF)], and γ (for RBF) using GridSearchCV under a fixed training/validation split. This controls bias-variance trade-offs and improves generalization to field variability. Model selection is driven by precision, recall, F1-score, and inspection of the confusion matrix.

Decision function:

$$f(x) = sign\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right) \tag{9}$$

Here x is the input feature vector, x_i (i = 1, ..., n) are the n training samples, $K(\ ,\ )$ is the kernel, α_i are Lagrange multipliers, y_i are labels, and b is the bias. Linear kernels favor speed and interpretability; RBF captures nonlinear lesion textures and backgrounds. We persist the fitted SVM together with PCA and scaling parameters for reproducible, real-time inference [17].

Random Forest ensembles many decorrelated decision trees, each trained on a bootstrap sample with random feature subsetting, yielding robust performance on heterogeneous leaf imagery. The input features for each decision tree consist of the same PCA-reduced and standardized raw pixel vectors from resized leaf images (224 × 224) used for SVM and k-NN. This ensures consistent representation across traditional classifiers and improves interpretability of feature importance scores [17].

Ensemble vote:

$$\hat{y}(x) = mode\left\{f_1(x),\ f_2(x),\ .....,\ f_t(x)\right\} \tag{10}$$

Here $f_t(x)$ is the class predicted by the t-th tree (t = 1, ..., T), and T is the number of trees. Majority voting reduces variance relative to single trees, improving stability under variable lighting, leaf pose, and lesion size. Trained models—and their preprocessing artifacts—are serialized for consistent deployment [17].

k-NN classifies a test sample by the majority class among its k closest training neighbors in feature space. Specifically, the features consist of raw pixel values from resized leaf images (224 × 224), which are then reduced to the top 20 principal components using PCA and standardized with StandardScaler. This representation is used to compute distances for k-NN classification, ensuring reproducibility and consistency with other traditional classifiers [17].

Distance (Euclidean):

$$d(u,\ v) = \sqrt{\sum_{j=1}^{m}(u_j - v_j)^2} \tag{11}$$

Here u and v are m-dimensional feature vectors, and u_j and v_j are their j-th components. The predicted label is the majority vote among the k nearest v to u. While inference scales with dataset size, k-NN provides a transparent baseline and strong performance when features are well-separated. We export the trained pipeline (PCA, scaler, classifier) for reliable, real-time classification [17].

Model evaluation

Early stopping based on the validation set is used to control the training of every model. The final results on the held-out test set are reported with the standard metrics given below:

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \tag{12}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{13}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{14}$$

$$F_1 = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right) \tag{15}$$

Because the same training criteria, data splits, and preprocessing are applied to every model, the contribution of each architecture can be examined in isolation. The results indicate that the custom CNN clearly surpasses every baseline, followed closely by ResNet-50 and VGG-16, whereas the traditional methods generalize less effectively by comparison [17].

Results

The performance of all models was evaluated on the held-out test set using the same preprocessing, data splits, and evaluation protocols. The comparison included deep learning models (ResNet-50, VGG-16, and the proposed custom CNN) and traditional machine learning classifiers (k-NN, SVM, and Random Forest). The training and validation behavior of deep learning models is illustrated in Figure 3, while confusion

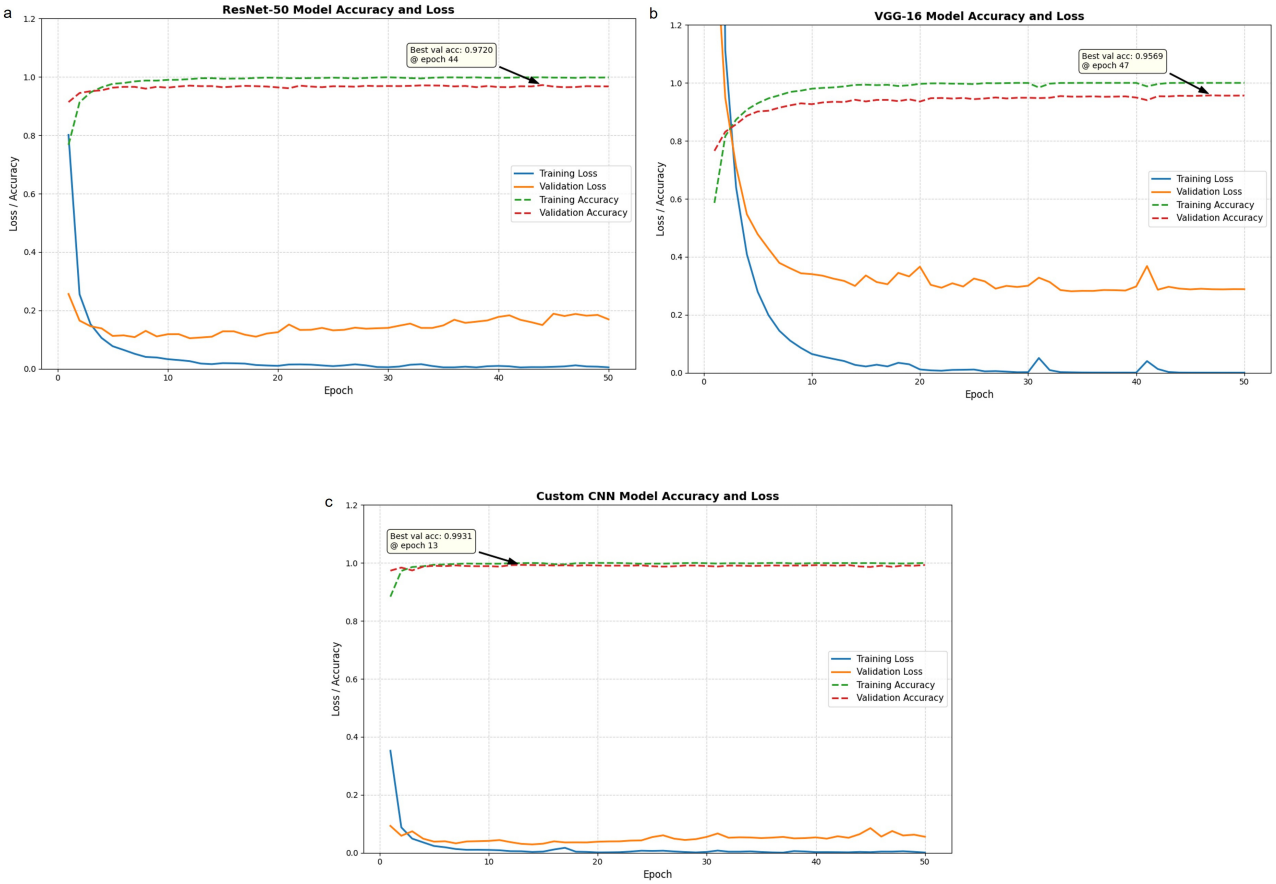


Figure 3. Learning curves (accuracy and loss) of the three deep learning models for multi-crop leaf disease recognition. (**a**) ResNet-50; (**b**) VGG-16; (**c**) custom CNN. CNN: convolutional neural network.

matrices for deep learning and traditional machine learning approaches are presented in Figures 4 and 5, respectively. The quantitative performance comparison is summarized in Tables 2 and 3.

Performance evaluation of deep learning models

As shown in Table 2, the proposed custom CNN delivered the best overall results, with 99.29% accuracy over all 14 classes, confirming that it recognizes leaf diseases effectively across multiple crops. We attribute this advantage to an architecture tailored to leaf images, which extracts multi-scale texture cues and fine morphological details of small lesions while distinguishing them from natural leaf structures. The model achieved perfect precision and recall for healthy rice leaves, while the F1-scores for tomato bacterial spot and potato late blight exceeded 0.98. Difficult categories with overlapping lesion characteristics, such as corn northern leaf blight and common rust, still reached approximately 97.04% accuracy, further confirming the robustness of the proposed approach. A learning-rate schedule, calibrated regularization, and balanced sampling kept training stable, while preprocessing and augmentation improved generalization to field-like conditions. Data augmentation included random rotations of up to ± 15°, left–right and up–down flips, and ± 10% changes in brightness, applied uniformly across all models. These design decisions account for the consistent advantage of the custom CNN over both the transfer-learning and traditional baselines, with dependable precision and recall for all crops and diseases. While the custom CNN performed exceptionally on the curated test set, these results should be interpreted cautiously. The dataset is balanced and curated, and the model may have encountered visually similar examples during training. Misclassifications were rare, mainly occurring between visually similar classes such as potato late blight versus early-stage healthy leaves and corn northern leaf blight versus common rust, highlighting the challenges of overlapping lesion morphology and variable backgrounds. Comparative evaluation showed that ResNet-50 and VGG-16 achieved 96.75% and 95.63% overall accuracy, respectively, as shown in Table 2, with minor sensitivity issues on low-contrast lesions. The traditional classifiers (*k*-NN, SVM, and Random Forest) were consistently weaker than the CNN, with overall accuracies between 89.71% and

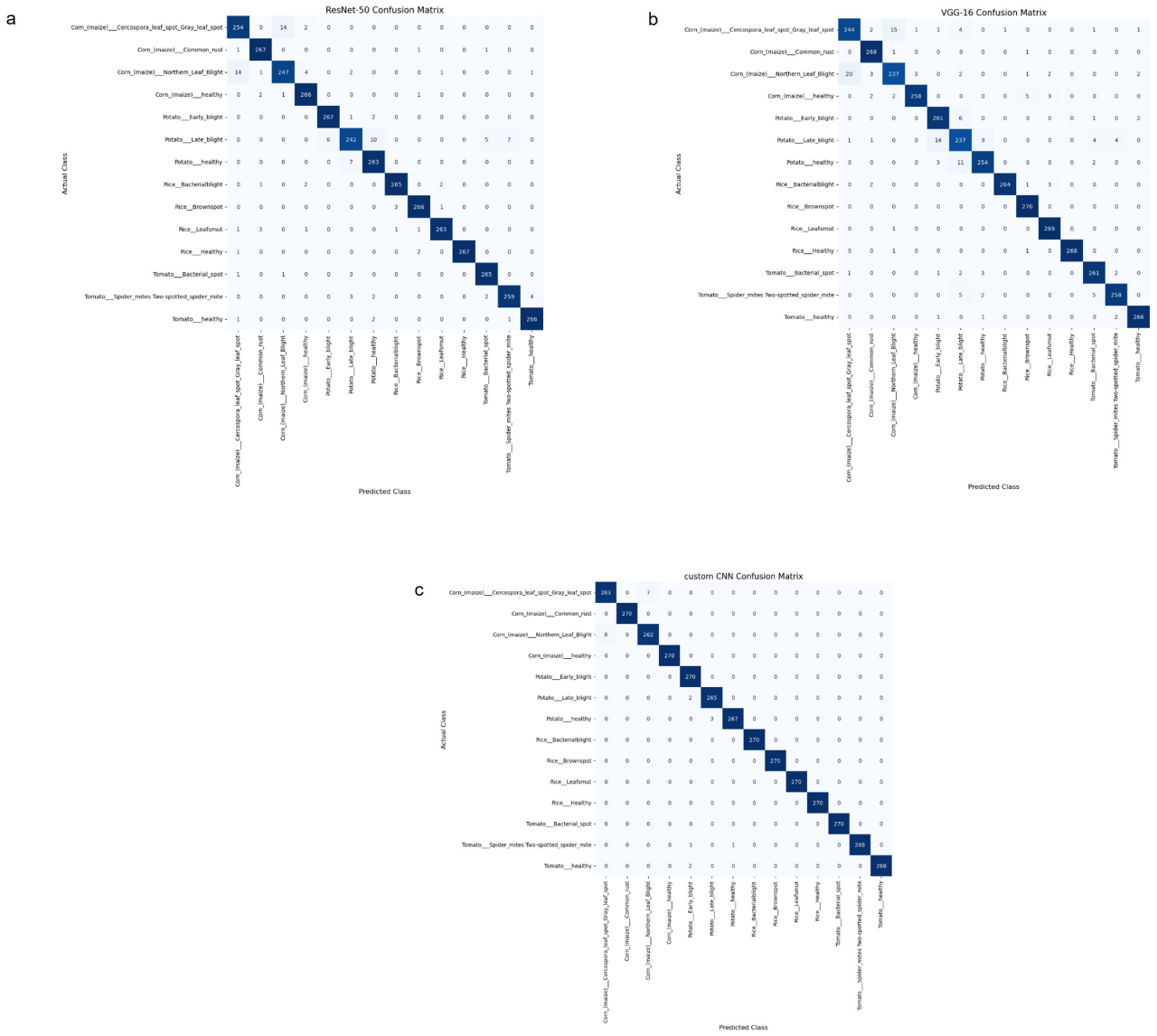


Figure 4. Confusion matrices produced by the three deep learning models for recognizing multi-crop leaf diseases. (**a**) ResNet-50; (**b**) VGG-16; (**c**) custom CNN.

90.87%, as shown in Table 3. These results underscore the superior feature extraction and robustness of the custom CNN under the given experimental conditions.

ResNet-50 demonstrated high accuracy of 96.75% and high accuracy on tomato bacterial spot with F1 score of 0.9761, precision of 0.9707, and potato early blight with an F1 score of 0.9834 as presented in Table 2. It had the lowest class of rice leaf smut in rice, which had F1 score of 0.9795, precision 0.9850 and its error analysis revealed it confused corn northern leaf blight where low-contrast lesions diminished sensitivity as we observe in Table 2 and Figure 4. These trends are in line with field imaging where mottled backgrounds and initial stages of symptoms squeeze lesion-leaf contrast. Nonetheless, these residual links provided effective convergence and stabilized optimization to produce robust cross-crop adaptability and competitive behavior among foliar morphologies that were different.

VGG-16 was closely behind with an accuracy of 95.63%, which is performing well on corn common rust with F1 score of 0.9781, precision of 0.9640 and rice brown spot with F1 score of 0.9836, but poorer on potato late blight with F1 of 0.8827, with recall of 0.8778, as illustrated in Table 2. The fine texture cues in the architecture presented by 3 × 3 uniform stacks were consistent and effective, but in cases of irregular or diffuse lesion boundaries. This is notwithstanding the limitation, which despite its weaknesses, allowed VGG-16 to have a balance in precision-recall profiles across most categories, which has provided a stable and reproducible baseline. Practically, they can be interpreted as improvements of the combination of

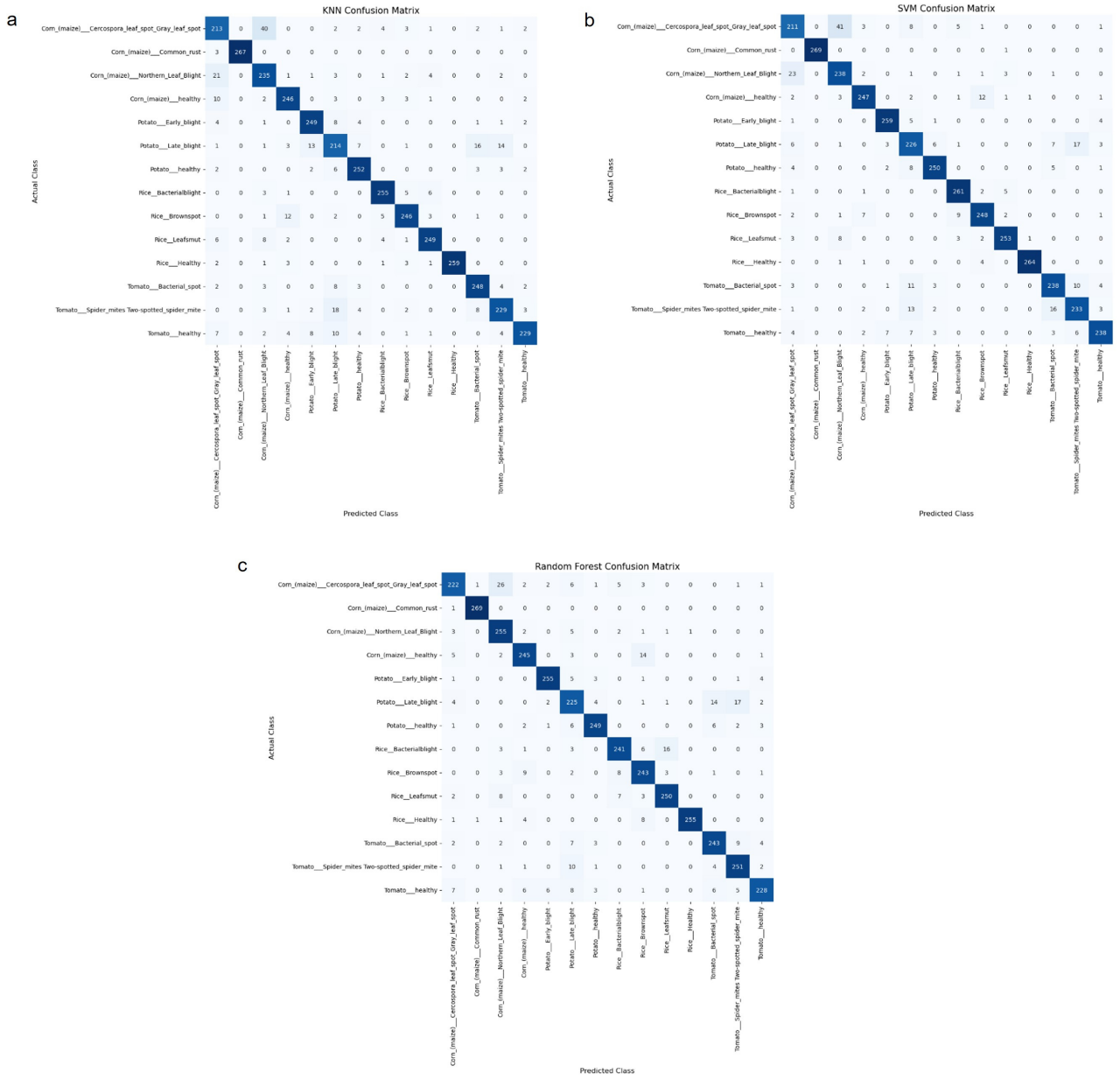


Figure 5. Confusion matrices produced by the three traditional machine learning models for recognizing multi-crop leaf diseases. (**a**) *k*-nearest neighbors (*k*-NN); (**b**) support vector machine (SVM); (**c**) Random Forest.

stronger augmentation and class-aware sampling with VGG-16 in terms of shape variability, and the usage of ResNet-50 when sensitivity in heterogeneous imaging conditions is prioritized as overviewed in Figure 4.

Performance evaluation of traditional ML models

k-NN attained an overall accuracy of 89.71% with an F1 score of 0.9944 on corn common rust and 0.8674 on tomato spider mite infestations where lesions are small, sparse and irregular, as evidenced in Table 3. Voting based on distance had difficulty separating classes with fine-grained, low-contrast patterns.

SVM was also found to have 90.87% accuracy with F1 of 0.9474 with rice bacterial blight separation, but the recall decreased in potato late blight 0.8370 recall as demonstrated in Table 3. The sensitivity of the margin-based classifier to heterogeneous lesion morphology and variation of backgrounds was also found to be a weak point in issue of sensitivity in field use.

Random Forest also achieved the accuracy of 90.77%, and it predicted the corn *Cercospora* leaf spot with an F1 of 0.8555 and tomato bacterial spot with F1 of 0.8934, as shown in Table 3. Figure 5 confusion matrices have more overlaps and misclassifications, reflecting the weaknesses of handcrafted features in the presence of high variability and the presence of minor inter-class differences.

Table 2. Comparison of deep learning models on multi-crop leaf diseases.

Class name	VGG-16			ResNet-50			Custom CNN		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Corn *Cercospora* leaf spot	0.9173	0.9037	0.9104	0.9304	0.9407	0.9355	0.9705	0.9741	0.9723
Corn common rust	0.9640	0.9926	0.9781	0.9745	0.9889	0.9816	1.0000	1.0000	1.0000
Corn northern leaf blight	0.9222	0.8778	0.8994	0.9392	0.9148	0.9268	0.9740	0.9704	0.9722
Corn healthy	0.9847	0.9556	0.9699	0.9673	0.9852	0.9761	1.0000	1.0000	1.0000
Potato early blight	0.9288	0.9667	0.9474	0.9780	0.9889	0.9834	0.9818	1.0000	0.9908
Potato late blight	0.8876	0.8778	0.8827	0.9380	0.8963	0.9167	0.9888	0.9815	0.9851
Potato healthy	0.9442	0.9407	0.9425	0.9427	0.9741	0.9581	0.9963	0.9889	0.9926
Rice bacterial blight	0.9962	0.9778	0.9869	0.9851	0.9815	0.9833	1.0000	1.0000	1.0000
Rice brown spot	0.9677	1.0000	0.9836	0.9815	0.9852	0.9834	1.0000	1.0000	1.0000
Rice leaf smut	0.9711	0.9963	0.9835	0.9850	0.9741	0.9795	1.0000	1.0000	1.0000
Rice healthy	1.0000	0.9926	0.9963	1.0000	0.9889	0.9944	1.0000	1.0000	1.0000
Tomato bacterial spot	0.9526	0.9667	0.9596	0.9707	0.9815	0.9761	1.0000	1.0000	1.0000
Tomato spider mites	0.9699	0.9556	0.9627	0.9700	0.9593	0.9646	0.9889	0.9926	0.9908
Tomato healthy	0.9815	0.9852	0.9834	0.9815	0.9852	0.9834	1.0000	0.9926	0.9963
Overall accuracy	0.9563			0.9675			0.9929		

CNN: convolutional neural network.

Table 3. Comparison of traditional machine learning models on multi-crop leaf diseases.

Class name	*k*-NN			SVM			Random Forest		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Corn *Cercospora* leaf spot	0.7860	0.7889	0.7874	0.8084	0.7815	0.7947	0.8916	0.8222	0.8555
Corn Common rust	1.0000	0.9889	0.9944	1.0000	0.9963	0.9981	0.9926	0.9963	0.9945
Corn northern leaf blight	0.7833	0.8704	0.8246	0.8123	0.8815	0.8455	0.8472	0.9444	0.8932
Corn healthy	0.9011	0.9111	0.9061	0.9321	0.9148	0.9234	0.9007	0.9074	0.9041
Potato early blight	0.9055	0.9222	0.9138	0.9522	0.9593	0.9557	0.9586	0.9444	0.9515
Potato late blight	0.7810	0.7926	0.7868	0.8043	0.8370	0.8203	0.8036	0.8333	0.8182
Potato healthy	0.9130	0.9333	0.9231	0.9434	0.9259	0.9346	0.9432	0.9222	0.9326
Rice bacterial blight	0.9341	0.9444	0.9392	0.9288	0.9667	0.9474	0.9163	0.8926	0.9043
Rice brown spot	0.9213	0.9111	0.9162	0.9185	0.9185	0.9185	0.8648	0.9000	0.8820
Rice leaf smut	0.9361	0.9222	0.9291	0.9547	0.9370	0.9458	0.9225	0.9259	0.9242
Rice healthy	1.0000	0.9593	0.9792	0.9925	0.9778	0.9851	0.9961	0.9444	0.9696
Tomato bacterial spot	0.8889	0.9185	0.9035	0.8815	0.8815	0.8815	0.8869	0.9000	0.8934
Tomato spider mites	0.8876	0.8481	0.8674	0.8759	0.8630	0.8694	0.8776	0.9296	0.9029
Tomato healthy	0.9463	0.8481	0.8945	0.9297	0.8815	0.9049	0.9268	0.8444	0.8837
Overall accuracy	0.8971			0.9087			0.9077		

k-NN: *k*-nearest neighbors; SVM: support vector machine.

Test sample recognition analysis

The proposed custom CNN demonstrated reliable classification performance on representative test samples, correctly identifying the actual and predicted classes of input leaf images, as illustrated in Figure 6. Across more than one hundred independent test samples, only one misclassification was observed, where a healthy leaf was incorrectly predicted as a diseased class. Specifically, a Tomato Healthy sample was misclassified as Tomato Bacterial Spot, highlighting the difficulty of distinguishing subtle early symptoms from naturally healthy leaf characteristics. Overall, the custom CNN achieved a strong balance between accuracy, robustness, and efficiency for multi-crop leaf disease recognition.

Web deployment demonstration

The interface will present every prediction with two clinician-facing components. Reasons, a summarization of the visual evidence and diagnostic indications such as lesion color, shape, distribution,

Actual class	Image Input	Predicted class	Result
Corn Gray Leaf Spot		Corn Gray Leaf Spot	✓ Correct
Corn Common Rust		Corn Common Rust	✓ Correct
Potato Early Blight		Potato Early Blight	✓ Correct
Potato Late Blight		Potato Late Blight	✓ Correct
Rice Brown Spot		Rice Brown Spot	✓ Correct
Rice Bacterial Blight		Rice Bacterial Blight	✓ Correct
Tomato Healthy		Tomato Bacterial Spot	✕ Wrong
Tomato Bacterial Spot		Tomato Bacterial Spot	✓ Correct

Figure 6. Recognition demonstration by the best-performing custom CNN model. CNN: convolutional neural network.

and context that caused the model to make that decision; and Solutions, which will transform that diagnosis into management steps, which are sanitation, cultural practices, resistant varieties, and where appropriate,

chemistry with mode-of-action rotation. The system assists in triage based on thresholds and coherent and defensible decisions in the field by combining calibrated probabilities with open and transparent reasons and guidance.

The trained model was deployed in a web application to demonstrate real-time inference. While a screenshot illustrates the interface and workflow, quantitative evaluation of live performance was not conducted; the system provides a framework for rapid leaf disease assessment in future practical use. The Multi-crop Leaf Disease Recognition system has an end-to-end workflow represented in Figure 7. The landing page (Figure 7a) provides brief instructions and a sample image, which directs the users to a one-session route, starting with uploading to advise. The ingestion view (Figure 7b) allows drag-and-drop/file browsing of JPEG/PNG images, which makes phone/laptop submission straightforward. The results view (Figure 7c) displays the best result of a calibrated confidence of the top class, the original image to be visually checked and Reasons and Solutions could be read by humans. The file can be cleared, and analyses can be re-run immediately. Standardized preprocessing and CNN inference provide consistency of inputs, constant latency, and predictable outputs under the hood. Developed to be used in checkpoints, the system is designed to be extended into clinics and used on the farm, so the model is both accurate and usable without handing over a lot of misclassified data to the decision-makers.

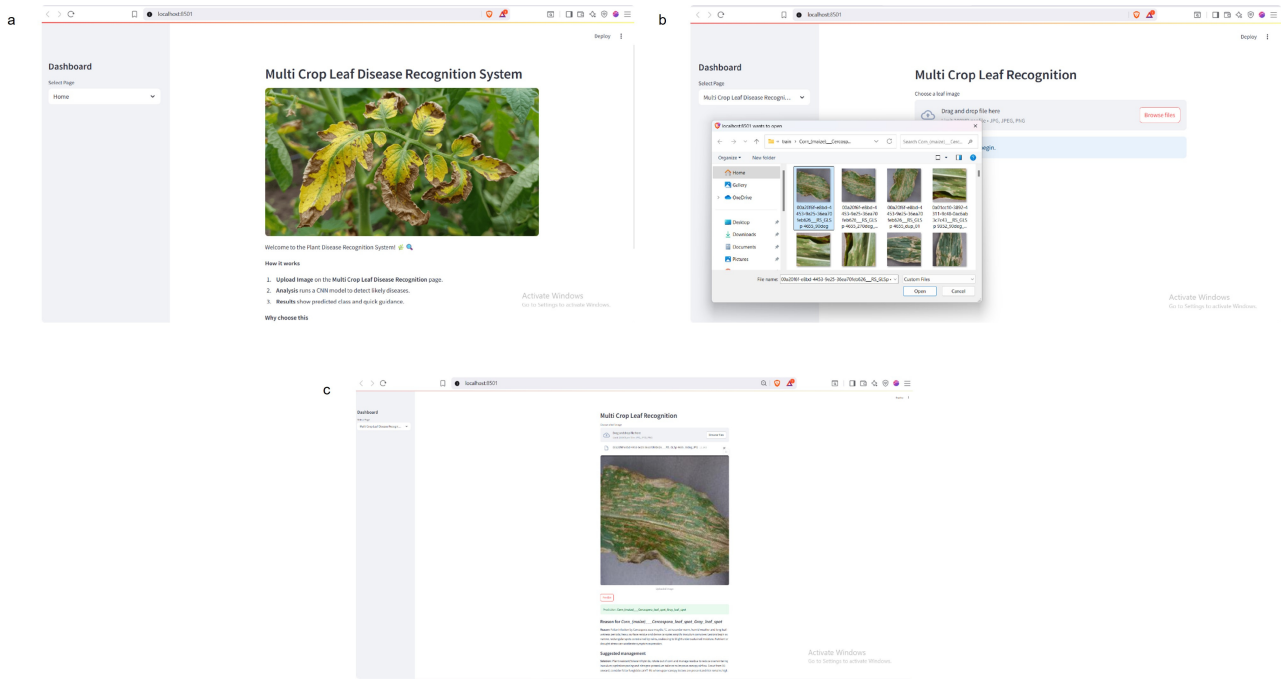


Figure 7. Demonstration of the multi-crop leaf disease recognition application. (a) Home screen; **(b)** image upload interface; **(c)** prediction output with the uploaded leaf image, predicted disease class, and recommended management guidance.

Discussion

The proposed framework demonstrates that a carefully designed custom CNN can achieve effective multi-crop leaf disease recognition while maintaining a simple and reproducible architecture. The balanced 14-class dataset, unified preprocessing strategy, and consistent evaluation protocol enabled a fair comparison among deep learning and traditional machine learning approaches. The proposed custom CNN achieved the highest classification performance with 99% accuracy on the test set, outperforming transfer-learning-based models such as ResNet-50 and VGG-16, as well as traditional classifiers including *k*-NN, SVM, and Random Forest. This improvement can be attributed to the model's ability to learn discriminative foliar patterns and extract multi-scale lesion features through its crop-specific convolutional design. Furthermore, the model maintained strong precision and recall across different crop-disease categories, demonstrating its reliability for automated plant disease recognition under the evaluated conditions.

Previous studies have also demonstrated that deep learning approaches work well for recognizing plant diseases, although, many existing methods focus on specific crops, limited disease categories, or specialized experimental settings. Pan et al. [4] designed a small YOLOv8-based detector with attention mechanisms for rice leaf diseases, achieving approximately 95% performance with a compact model structure. Saleem et al. [5] combined few-shot learning with feature fusion in a NASNetMobile-based hybrid model for classifying diseases of multiple crops, achieving competitive classification accuracy. Bouacida et al. [7] presented a small Inception-style CNN that concentrates on lesion areas and demonstrated strong performance on both controlled and unseen crop datasets. Rajagopal et al. [14] explored a classical image-processing pipeline combining segmentation, handcrafted features, and multi-kernel SVM for leaf disease classification. Taken together, these studies show that deep learning and classical machine learning approaches are both effective; however, their evaluation settings, crop coverage, and dataset compositions differ significantly from the unified multi-crop evaluation performed in this study.

As summarized in Table 4, the proposed framework provides a comprehensive benchmark by simultaneously considering four major crops and fourteen healthy/diseased classes using a balanced dataset and identical training protocols. Unlike approaches designed for individual crops or specific disease detection tasks, this study evaluates multiple deep learning and traditional machine learning models under the same experimental conditions. The proposed custom CNN achieves competitive performance with a relatively lightweight architecture while avoiding the complexity associated with large-scale detection networks or ensemble solutions. The standardized experimental pipeline, including fixed data splits, consistent preprocessing, and identical evaluation metrics, improves reproducibility and enables a fair comparison of different learning approaches for multi-crop disease recognition.

Table 4. Comparison of different methods for crop leaf disease recognition.

Work done	Problem solved	Key characteristics	Model name	Dataset size	Classes	Accuracy
This work	Multi-crop disease classification (corn, potato, rice, tomato)	Balanced dataset; foliar-specific CNN; class-balanced learning; unified training	Custom CNN	21,000 images	14	99%
Pan et al. [4]	Rice leaf disease detection (fresh/half/spoiled)	YOLOv8-based tiny detector; SENet attention	SSD-YOLO	3,000 images	3	95.3% (mAP)
Saleem et al. [5]	Multi-crop Disease classification (potato, tomato, mango)	NASNetMobile backbone; vegetation index preprocessing; feature fusion; few-shot learning	AgirLeafNet	~17,000 images	~15	98% (3-class accuracy)
Bouacida et al. [7]	Cross-crop classification (generalized lesions)	Compact inception-like CNN; focuses on lesion regions; cross-crop generalization	Lesion-Focused CNN	54,000 images	38	94.0% (lab); 97.1% (unseen)
Rajagopal et al. [14]	Classical pipeline (segmentation + SVM)	Fuzzy C-means segmentation; handcrafted vein/texture features; multi-kernel SVM	LK-PSVM	Small lab set	3	96.1%

CNN: convolutional neural network; SVM: support vector machine.

Despite achieving strong classification performance, several limitations should be acknowledged. The current evaluation was conducted using a balanced and curated dataset, which may not fully represent the variability of real agricultural environments. Factors such as complex backgrounds, illumination changes, leaf orientation, and varying disease severity may affect model performance during practical deployment. Therefore, additional validation using large-scale field-collected images is necessary to confirm real-world generalization. Future research should investigate domain adaptation techniques, lightweight edge deployment, and continuous learning strategies to improve the robustness and practical applicability of automated crop disease recognition systems.

In conclusion, we presented a robust and generalizable approach to multi-crop leaf disease recognition, aiming to bridge the gap between controlled research evaluation and practical agricultural diagnostic requirements. Although the curated 21,000-image dataset draws from existing open-source sources, its

balanced multi-crop structure and standardized training pipeline enable rigorous and reproducible evaluation across multiple models and crop types. In this context, our foliar-specific CNN consistently outperformed state-of-the-art deep models (e.g., ResNet-50, VGG-16) and traditional machine learning models *k*-NN, SVM and Random Forest, delivering near-perfect classification accuracy across all crop and disease categories. While the architecture follows standard CNN design principles, its domain-specific optimization for leaf lesion recognition contributed to high performance across multiple crops and disease classes. This high performance underscores the value of a tailored architecture and disciplined training strategy in enabling reliable detection on curated datasets, without implying full validation in independent field conditions. Collectively, the work sets a new performance benchmark for multi-crop disease identification. While the results demonstrate strong performance on curated datasets, the model's effectiveness in independent field scenarios remains to be validated. The deployment in a web application illustrates its potential for on-site disease monitoring and rapid diagnostic support, providing a framework for future field trials and practical applications in precision agriculture. Future work will include on-farm trials and validation on additional crops and varied field conditions to confirm generalization beyond the curated dataset. We also aim to conduct on-farm trials and develop mobile or edge-based applications to further validate and refine the model's performance under diverse real-world conditions.

Abbreviations

CNN: convolutional neural network

k-NN: *k*-nearest neighbors

PCA: principal component analysis

RBF: radial basis function

SVM: support vector machine

ViT: vision transformer

Declarations

Author contributions

ZWB: Conceptualization, Methodology, Software, Investigation, Data curation, Formal analysis, Visualization, Resources, Writing—original draft, Writing—review & editing. ZBMA: Writing—original draft. MH: Supervision, Conceptualization, Resources, Writing—review & editing. MTH: Conceptualization, Methodology, Supervision, Resources, Project administration, Writing—review & editing. All authors read and approved the submitted version.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

All images used in this study were obtained from publicly available repositories, primarily the rice leaf diseases dataset (Antony and Prasanth [1]), the seasonal corn leaf disease dataset (Ahmad [2]), and the

potatof leaf disease dataset in uncontrolled environments (Shabrina et al. [3]), all hosted on Mendeley Data. These images were re-curated into a balanced 14-class dataset of 21,000 images covering corn, potato, rice, and tomato, with fixed training, validation, and test splits, as described in the Materials and methods section. No new raw data were generated in this study, and all source data can be accessed directly from the repositories cited above.

Funding

Not applicable.

Copyright

© The Author(s) 2026.

Publisher's note

Open Exploration maintains a neutral stance on jurisdictional claims in published institutional affiliations and maps. All opinions expressed in this article are the personal views of the author(s) and do not represent the stance of the editorial team or the publisher.

References

1. Antony L, Prasanth L. Rice leaf diseases dataset [Internet]. [cited 2025 Jul 21]. Available from: https://data.mendeley.com/datasets/dwtn3c6w6p/1
2. Ahmad H. Seasonal Corn Leaf Disease Dataset: A Multi-Year Collection for Robust Analysis [Internet]. [cited 2025 Jul 21]. Available from: https://doi.org/10.17632/vy629dngm8.1
3. Shabrina HN, Indarti S, Maharani R, Kristiyanti DA, Irmawati I, Prastomo K. Potato leaf disease dataset in uncontrolled environment [Internet]. [cited 2025 Jul 21]. Available from: https://doi.org/10.17632/ptz377bwb8.1
4. Pan C, Wang S, Wang Y, Liu C. SSD-YOLO: a lightweight network for rice leaf disease detection. Front Plant Sci. 2025;16:1643096. [DOI] [PubMed] [PMC]
5. Saleem S, Sharif MI, Sharif MI, Sajid MZ, Marinello F. Comparison of Deep Learning Models for Multi-Crop Leaf Disease Detection with Enhanced Vegetative Feature Isolation and Definition of a New Hybrid Architecture. Agronomy. 2024;14:2230. [DOI]
6. Tunio MH, Li JP, Zeng X, Ahmed A, Shah SA, Shaikh H, et al. Advancing plant disease classification: A robust and generalized approach with transformer-fused convolution and Wasserstein domain adaptation. Comput Electron Agric. 2024;227:109574. [DOI]
7. Bouacida I, Farou B, Djakhdjakha L, Seridi H, Kurulay M. Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves. Inf Process Agric. 2025;12:54–67. [DOI]
8. Islam M, Adil AA, Talukder A, Ahamed KU, Uddin A, Hasan K, et al. DeepCrop: Deep learning-based crop disease prediction with web application. J Agric Food Res. 2023;14:100764. [DOI]
9. Khan AT, Jensen SM, Khan AR, Li S. Plant disease detection model for edge computing devices. Front Plant Sci. 2023;14:1308528. [DOI] [PubMed] [PMC]
10. Feng J, Ong WE, Teh WC, Zhang R. Enhanced Crop Disease Detection With EfficientNet Convolutional Group-Wise Transformer. IEEE Access. 2024;12:44147–62. [DOI]
11. De Silva M, Brown D. Multispectral Plant Disease Detection with Vision Transformer–Convolutional Neural Network Hybrid Approaches. Sensors. 2023;23:8531. [DOI] [PubMed] [PMC]
12. Hemalatha S, Jayachandran JJB. A Multitask Learning-Based Vision Transformer for Plant Disease Localization and Classification. Int J Comput Intell Syst. 2024;17:188. [DOI]
13. Baek ET. Attention Score-Based Multi-Vision Transformer Technique for Plant Disease Classification. Sensors. 2025;25:270. [DOI] [PubMed] [PMC]

14. Rajagopal M, Kayikci S, Abbas M, Sivasakthivel R. A novel technique for leaf disease classification using Legion Kernels with parallel support vector machine (LK-PSVM) and fuzzy C means image segmentation. Heliyon. 2024;10:e32707. [DOI] [PubMed] [PMC]
15. Ngugi HN, Akinyelu AA, Ezugwu AE. Machine Learning and Deep Learning for Crop Disease Diagnosis: Performance Analysis and Review. Agronomy. 2024;14:3001. [DOI]
16. Demilie WB. Plant disease detection and classification techniques: a comparative study of the performances. J Big Data. 2024;11:5. [DOI]
17. Bhuiyan ZW, Limon MH, Rana MR, Akthar SR, Habib MT, Hasan M. A Comparative Study of Deep Learning Models for Plant Leaf and Disease Recognition in Jackfruit, Rice, and Rose. In: Iwendi C, Boulouard Z, Kryvinska N, editors. Proceedings of the 4th International Conference on Advances in Communication Technology and Computer Engineering (ICACTCE'24); 2025 Aug 30; Cham, Switzerland. Cham: Springer; 2024. pp. 460–71. [DOI]